Euro Area Inflation and Unemployment: Phillips Curve and Forecasting*

Emrehan Aktuğ
Sabancı University

Saroj Bhattarai
UT Austin

Ayşe Dur
NC State

Abstract

We study Euro Area inflation and unemployment dynamics by addressing the identification challenges inherent to estimating the Phillips curve and providing new frameworks and results for inflation forecasting. Phillips curve slope estimation is challenging due to the simultaneity between monetary policy and economic conditions that biases OLS estimates downward. To overcome this challenge, we adopt a panel local projection IV framework. We estimate the cumulative impacts of monetary policy shocks on inflation and unemployment over a given horizon, where we use U.S. monetary policy shocks as IVs to ensure a causal interpretation of our results. We show that the policy trade-off identified through this panel IV estimation of the Phillips curve relationship is not just qualitatively, but also quantitatively, different from OLS estimation. We find no evidence that this trade-off has changed recently in the post-pandemic period. Finally, we provide novel tools for inflation forecasting based on the panel Phillips curve model. We document that panel data models yield accurate one-year ahead forecasts for 17 member states in the 2009-2023 period and clearly outperform the commonly used, naive single-country models.

JEL classification: E31, C33, C36, E37

Keywords: Euro Area; Inflation; Unemployment; Phillips Curve; Panel Local Projection IV; Inflation Forecasting; Panel Models

*We thank Barbara Rossi, Jonathan Wright, Regis Barnichon, Michael Bauer, and Tatevik Sekhposyan for very helpful comments. First version: April 2025. This version: June 2025. Aktug: emrehan.aktug@sabanciuniv.edu; Bhattarai: saroj.bhattarai@austin.utexas.edu; Dur: ayse.z.dur@ncsu.edu.

1 Introduction

> *"The challenge we face was well-captured by the philosopher Søren Kierkegaard, who said that 'life can only be understood backwards; but it must be lived forwards.' Since our policies operate with lags, we cannot wait for the parameters of this new environment to become entirely clear before we act. We have to form a view of the future and act in a forward-looking way..."*
>
> — Christine Lagarde, Jackson Hole, 2023

Understanding the nexus between inflation and real economic activity is central to the objectives of monetary policy. The Phillips curve traditionally captures this short-term relationship, linking inflation to real economic slack measures such as unemployment, and thereby provides a measure of monetary policy trade-off. However, uncovering this relationship in the data is notoriously hard due to identification difficulties, and this has led some to question whether the Phillips curve exists at all. In this paper, we pursue two main research questions. First, we reexamine the inflation-unemployment trade-off in the Euro Area, addressing long-standing identification challenges in a panel estimation setting. Second, we propose novel panel data models for forecasting inflation. We provide robust evidence—both in-sample and out-of-sample—on the connections between monthly Euro Area inflation and unemployment over the 1999-2023 period, covering 17 Euro Area member countries. This period poses significant challenges for researchers, as it includes both the prolonged low-inflation environment preceding the pandemic and the sharp surge in inflation that followed. We show that our empirical approaches are very useful in understanding inflation dynamics in this period.

While the Phillips curve is central to quantifying the employment cost of reducing inflation, a major pitfall for this exercise is simultaneity bias: both inflation and unemployment are jointly determined by the interaction of the Phillips curve and the central bank's policy response. Consequently, simple regressions of inflation on unemployment often underestimate the strength of the relationship. In fact, one might even find the wrong sign: The data may suggest a positive relationship between unemployment and inflation, which arises precisely because monetary policy seeks to lower unemployment (by loosening) when inflation is below the target and raise unemployment (via tightening) when inflation is above the target.[1] In addition, the responses of unemployment and inflation

[1] For example, in the textbook sticky price model, under optimal monetary policy with cost push shocks, the targeting rule features a negative relationship between inflation and output gap. Thus, even though the model features a traditional Phillips curve with a positive slope between inflation and output gap, a simple OLS regression would yield a negative coefficient.

to monetary policy evolve dynamically, possibly following different lags, making it even more difficult to pin down their true contemporaneous relationship, as Christine Lagarde also pointed out in her Jackson Hole speech in 2023.

Our identification and estimation strategy features two key aspects. First, we follow the remedies prescribed in the literature by using data from a currency union, employing an instrumental variable (IV) to isolate demand-side variation, and controlling for global cost-push shocks. Second, to capture the dynamic and differential responses of unemployment and inflation to common external policy shocks, we consider a Phillips multiplier model in the spirit of Barnichon and Mesters (2021), by adopting a *panel local projection* (LP).[2] Furthermore, adopting a panel model enables the use of a number of fixed effects to control for inflation expectations common to the currency union, country-specific trends, and global supply shocks, where limited data availability or data imperfections pose additional empirical challenges.[3]

Our panel LP approach captures the inflation-unemployment trade-off in a forward-looking sense, based on the relationship between cumulative inflation and cumulative unemployment from the present up to h-periods ahead. This multiplier is also the inverse of the traditional *sacrifice ratio*, which helps policymakers quantify the economic cost of disinflation. Accordingly, the sacrifice ratio measures the cumulative increase in the unemployment rate over a horizon due to a one percentage point deflation over the same period. Our panel LP-IV strategy involves a high-frequency identified, foreign, monetary policy shock as an instrument for (cumulative) unemployment. Specifically, we use US monetary policy shocks by Bauer and Swanson (2023) interacted by country fixed effects.

We document highly robust Phillips curve slope estimates across both headline and core inflation measures, for 17 Euro Area members and extending into the post-pandemic period. Our IV estimation results reveal a steeper slope estimate for the Phillips curve compared to OLS estimation in the sample period, verifying the downward bias inherent to OLS estimation of Phillips curve slopes. Finally, our Phillips curve multiplier estimates are stable over time, and we do not find evidence for a change in the multiplier while including data from the post-COVID-19 sample. This stability in estimates is still robust after controlling for the relative price of tradables and nontradables in our estimation.[4]

[2] In Appendix E, we show how for demand shocks, the standard textbook sticky price model augmented with internal sources of dynamics leads to a dynamic Phillips multiplier, analogous to our empirical model and estimates.

[3] Inflation expectations that play a key role in New Keynesian models alone pose another challenge for identification. The comprehensive survey by Mavroeidis et al. (2014) concludes that the literature might have already reached a limit with standard macroeconomic data.

[4] For a recent theoretical analysis of how the relative price of a tradable sector can affect headline inflation dynamics, through the role of input-output linkages and heterogeneous price stickiness, see Afrouzi et al.

Our paper next makes a novel contribution by evaluating the forecasting performance of the panel Phillips curve model. Specifically, to forecast 12-month ahead inflation, we consider current-period unemployment and lags of inflation.[5] We leverage this type of panel structure in alternative models informed by the structural models considered in in-sample analysis.

The key takeaways from the forecasting exercise are as follows. First, the panel Phillips curve model consistently outperforms a strong benchmark, the AR process of inflation in both pre- and post-COVID periods, EA12 and EA17 country groups, and both non-tradables and headline inflation measures. Second, the absolute RMSEs of these naive benchmarks, as well as a new framework—the panel AR(12) model—have generally increased in the post-COVID period, suggesting greater forecast uncertainty. Third, the panel AR(12) model emerges as a particularly strong forecasting tool and may serve as a new benchmark for inflation forecasting. In several robustness checks, this model in fact frequently outperforms the panel Phillips curve approach.

A large literature has examined the Phillips curve and reached mixed conclusions about its slope, highlighting an unstable relationship over time and across countries, as well as different macroeconomic data. Moreover, there is also no consensus regarding how the Phillips curve must be specified. Mavroeidis et al. (2014), McLeay and Tenreyro (2020), and Furlanetto and Lepetit (2024) explain the general difficulties of estimating the Phillips curve, and conclude that the literature has reached a limit with traditional approaches.

Among these, McLeay and Tenreyro (2020) argue that traditional time-series approaches, especially within single-country settings, are particularly susceptible to these problems because monetary policy decisions often mute the variation in inflation and aggregate slack needed for precise estimation. As a remedy, McLeay and Tenreyro (2020) propose exploiting cross-country variation—particularly within currency unions such as the Euro Area. Their key insight is that, while all member countries share a common monetary policy set by the ECB, they display heterogeneous economic conditions (e.g., levels of slack). This divergence allows researchers to better isolate the inflation-slack relationship by removing the confounding influence of country-specific monetary policy responses, thereby addressing the simultaneity problem and yielding more reliable estimates of the Phillips curve.

In addition, leveraging disaggregated data—at the regional, state, or industry level—

(2024).

[5] While we retain the panel structure and rely on the inflation-unemployment trade-off studied in the panel local projection model, we abandon the *cumulative* measures to tailor this model for forecasting purposes.

can have additional advantages since this type of data offer richer variation that enhances the precision of slope estimates in the Phillips curve. Recent empirical studies have followed these insights, applying it to Euro Area panels (e.g., Wellmann (2023), closely following Hazell et al. (2022), and Eser et al. (2020)), regional U.S. panels (e.g., Hazell et al. (2022); McLeay and Tenreyro (2020); Barnichon and Shapiro (2024); Fitzgerald et al. (2024)), panels of both regional US and Euro Area (Smith et al. (2025)); and panels of advanced economies (e.g., Jordà and Nechio (2018); Gabriel (2023)). Several of these studies go a step further by employing instrumental variables approaches to achieve causal identification of the Phillips curve slope.

The Phillips multiplier framework proposed by Barnichon and Mesters (2021) has gained attention for its flexible approach to modeling the Phillips curve. Using this specification, they documented historical changes in the slope of the relationship between inflation and slack, applying the model to U.S. and U.K. data. Several studies have since estimated variants of this model, including Gabriel (2023), Eser et al. (2020), and Inoue et al. (2024). Gabriel (2023) focused on a wage inflation Phillips curve for a panel of advanced economies, emphasizing historical instability and identification challenges. Eser et al. (2020) estimated the model using a Euro Area panel,[6] while Inoue et al. (2024) adopted a time-varying parameter IV approach to document changes in the Phillips curve with US time series within this framework. Our paper differs from these papers in jointly considering a 17-country Euro Area panel data approach *and* IV estimation, and by adopting a strong instrument supported by first and second stage results.

By mitigating identification challenges and uncovering more accurate slope estimates in the Phillips curve, we next turn to the question of out-of-sample performance in our paper. To our knowledge, this paper is the first to apply panel-based Phillips curve models for inflation forecasting. This is the second major dimension distinguishing our work from the previous literature.

In the next two sections, we present our empirical methodologies, data, and econometric and forecasting results.

2 Empirical Methodology and Data

2.1 Inflation-Unemployment Trade-off in a Currency Union

In a currency union, monetary policy is conducted centrally, implying that all member states are subject to the same monetary policy. Building on this, we extend the single-

[6] While they also explored an IV strategy, it was implemented only on union-wide data.

country Phillips multiplier framework of Barnichon and Mesters (2021) to a panel setting, where a uniform monetary policy shock, in our context a foreign monetary policy shock, generates common impulse responses for inflation and unemployment across countries. This approach is consistent with the theoretical model in Hazell et al. (2022), where, under the assumptions made on preferences and technology across countries, the slope of the regional Phillips curve equals that of the aggregate Phillips curve, suggesting a unified inflation-unemployment trade-off across member states.[7]

In this setting, the central bank's trade-off at horizon h is characterized by a single slope, estimated using common responses to an external (U.S.) monetary policy shock. We view this shock as a demand shock for the currency union countries. Specifically, we define the Phillips multiplier β_h as:

$$\beta_h = \frac{\bar{g}_h^{\pi}}{\bar{g}_h^{u}}, \quad h \geq 0 \tag{1}$$

where $\bar{g}_h^{\pi} = \frac{1}{h}\sum_{j=0}^{h} g_{n,j}^{\pi}$ and $\bar{g}_h^{u} = \frac{1}{h}\sum_{j=0}^{h} g_{n,j}^{u}$ denote the common, average impulse responses of inflation and unemployment, respectively, in response to a one-unit policy shock $\varepsilon_t^i = 1$.

2.2 A Panel Phillips Multiplier Model

We consider a panel with $n = 1, \ldots, N$ countries and $t = 1, \ldots, T$ periods. Let $\pi_{n,t}$ and $u_{n,t}$ denote inflation and unemployment in country n at time t. The Phillips multiplier, β_h, measures the cumulative inflation-unemployment trade-off over a horizon of h periods in the following panel:

$$\sum_{j=0}^{h} \pi_{n,t+j} = \beta_h \sum_{j=0}^{h} u_{n,t+j} + W_{n,t}'\gamma_h^{\pi} + \alpha_{nh} + \gamma_{th} + \varepsilon_{n,t+h}^{\pi} \quad \text{for } h = 0, 1, \ldots, H \tag{2}$$

where $W_{n,t}$ denotes control variables, and α_{nh}, γ_{th} represent country and time fixed effects, respectively. Time fixed effects can account for union-wide, long-term inflation expectations and macroeconomic shocks, consistent with Hazell et al. (2022). The set of controls in $W_{n,t}$ includes inflation lags (we consider 12 lags in monthly data) and country trends, following Barnichon and Shapiro (2024). We run a separate regression to estimate coefficients that vary with h.

[7] See Equations (12) and (13) in Hazell et al. (2022).

Identification Strategy. Identification of the Phillips multiplier in equation (2) requires a valid instrument for the cumulative unemployment term. To pursue an instrumental variable (IV) approach, we assume all countries are exposed to a common, exogenous, and external monetary policy shock, ε_t^i, for which we use a proxy instrument:

$$\xi_t = \mu\varepsilon_t^i + \nu_t, \tag{3}$$

where ν_t is classical measurement error.

To exploit cross-country variation, it is possible to construct an instrument, interacting a common external shock ξ_t (e.g. a common monetary policy shock) with a country-specific exposure measure $\mu_{n,t}$ (e.g. trade weights):

$$\text{MShock}_{n,t} = \mu_{n,t} \cdot \xi_t = \mu_{n,t}(\mu\varepsilon_t^i + \nu_t). \tag{4}$$

which yields *one* country-time level instrument $MShock_{n,t}$. As is common in the empirical literature, one can alternatively consider *multiple* instruments. In this case, we can introduce a set of country dummies z_n, taking a value equal to one for a given country and zero otherwise, for $n = 1, ..., N$ and again, interacting them with a common, exogenous variable, ξ_t^i. This leads to a total of N country-time level instruments, where, again the shift-share structure follows the logic of Bartik instruments, widely used in regional economics and macroeconomic panel settings to generate cross-sectional heterogeneity from a common shock. The common shock used in this setting can be an exogenous monetary policy shock.

Estimation Strategy. Our estimation framework is given in equation (2). With a *single* instrument, the estimation strategy can be relatively straightforward. We note that using local projections based on Jordà (2005); Jordà and Taylor (2025), one can also specify the impulse responses of inflation and unemployment separately in a panel of member states. An early panel application for OECD countries in the context of fiscal multipliers was adopted by Auerbach and Gorodnichenko (2012). This modeling strategy allows an econometrician to estimate horizon-h response equations for inflation and unemployment, respectively, as described below:

$$\pi_{n,t+h} = MShock_{n,t}\beta_h^\pi + \pi_{n,t-1}\rho_h^\pi + W'_{n,t}\theta_h^\pi + \alpha_{nh}^\pi + \gamma_{th}^\pi + \varepsilon_{n,t+h}^\pi \tag{5}$$

$$u_{n,t+h} = MShock_{n,t}\beta_h^u + u_{n,t-1}\rho_h^u + W'_{n,t}\theta_h^u + \alpha_{nh}^u + \gamma_{th}^u + \varepsilon_{n,t+h}^u. \quad (6)$$

In the panel local projections (5)-(6), $\pi_{n,t+h}$ and $u_{n,t+h}$ denote the h-step-ahead inflation and unemployment outcomes for country n at time $t+h$, respectively. The main regressor of interest, $MShock_{n,t}$, captures identified U.S. monetary policy shocks interacted with a country-time variable such as trade weights. These are treated as exogenous drivers of macroeconomic dynamics and are assumed to be common across countries at time t. Still, their effects may vary with country characteristics and evolve over horizons $h = 0, 1, \ldots, H$. The terms $\pi_{n,t-1}$ and $u_{n,t-1}$ denote the lagged inflation and unemployment levels, included to account for serial correlation and persistence in the respective variables.

The vector $W_{n,t}$ includes a set of control variables such as inflation lags and country-specific trends (e.g., yearly linear trends), which help capture evolving macroeconomic conditions. Country fixed effects α_{nh}^{π} and α_{nh}^{u} control for time-invariant unobserved heterogeneity across countries, while time fixed effects γ_{th}^{π} and γ_{th}^{u} account for inflation global shocks or common time trends affecting all countries. Since equations (5)-(6) are two separate regressions, time fixed effects may absorb different types of global shocks in each regression.[8] The estimated values of $\{\beta_h^{\pi}\}_{h=0}^{H}$ and $\{\beta_h^{u}\}_{h=0}^{H}$ are the impulse responses of each outcome to a monetary policy shock.

A three-step approach to estimate a multiplier is discussed in Ramey and Zubairy (2018) in the context of government spending multipliers. Accordingly, (i) one can estimate equation (5) for inflation and find the impulse responses β_h^{π} for $t = 0, 1..., h$, (ii) then estimate equation (6) for unemployment and find the impulse responses β_h^{u} from $h = 0, 1..., H$, (iii) and compute the multiplier as the ratio of the impulse responses from steps (i)-(ii), for $h = 0, 1..., H$.

The one-step IV approach in equation (2), where $MShock_{n,t}$ is the instrument for $\sum_{j=0}^{h} u_{n,t+j}$, yields an identical estimate to this three-step approach. For practical reasons, the commonly adopted strategy to estimate multipliers in Ramey and Zubairy (2018) and Barnichon and Mesters (2021) is also the one-step IV approach, which we follow in this paper.[9]

In the case of *multiple* instruments, which is the approach we implement in this paper,

[8] Forbes (2019), among others, finds that the Phillips curve has flattened over the past decades in a panel of advanced and emerging economies, primarily due to global factors such as global slack and commodity prices, with headline inflation particularly influenced by these forces. Our specification, therefore, captures such channels.

[9] The three-step approach poses several challenges in practice, as extensively discussed by Jordà and Taylor (2025). Estimating the standard errors of the ratio of two coefficients (i.e. the standard error of the implied multiplier) necessitates using approximation techniques, adding complexity to the procedure. Moreover, the ratio's denominator (specifically, β_h^u) may approach zero in some cases, potentially introducing instability and bias into the estimated standard errors.

we directly estimate equation (2) using an IV strategy. Specifically, we instrument for the cumulative unemployment term using N country-specific instruments constructed by interacting Bauer and Swanson (2023) U.S. monetary policy shocks with country dummies. This setup introduces first-stage heterogeneity, since each country responds differently to the common shock in the first-stage regression. As a result, it is no longer valid to assume common impulse responses across countries for inflation and unemployment as in equations (5)-(6). However, the second-stage regression still imposes a common slope on the inflation-unemployment relationship, which allows us to estimate a common Phillips multiplier and the associated sacrifice ratio. This is analogous to recent applications of Bartik-style IV strategies in regional macroeconomics (see, for example, Nakamura and Steinsson (2014) where cross-sectional heterogeneity in exposure is leveraged to identify a common structural parameter.

Finally, since equation (5) can have direct implications for inflation forecasting, we consider the N-country version of it in the out-of-sample forecasting exercises in section 3.2 in this paper.

2.3 Euro Area Panel Data

The sample period and country coverage are determined by data availability and quality. We use monthly data from 17 Euro Area members over the period 1999-2023. Our measure of slack is the monthly unemployment rate. Inflation is constructed at a monthly frequency as year-over-year changes in both core and headline price indexes, eliminating the need for seasonal adjustment. Core inflation is based on the nontradable goods sector. We also construct a relative price of tradable to the nontradable sector, and use it for some additional analysis.[10]

The inclusion of inflation lags in our models implies that the effective sample starts in 2000. We conduct econometric analyses for two time periods: a pre-pandemic subsample (2000m1-2019m12) and the full sample (2000m1-2023m12). We examine both the original Euro Area members "EA12", which includes Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, and the expanded "EA17", which includes the late joiners Cyprus, Estonia, Malta, Slovakia, and Slovenia.[11] Our main analysis focuses on the EA17 sample.

[10] For further details on how tradable and non-tradable prices are calculated, please see Appendix-D.

[11] Latvia and Lithuania are excluded due to their late accession and macroeconomic fundamentals that differ markedly from the rest of the sample.

2.4 Euro Area Phillips Correlation

In the Euro Area, the inflation-unemployment trade-off has become increasingly salient in the wake of COVID-19 pandemic. As the literature has pointed out (see, e.g. Benigno et al. (2023) and Hobijn et al. (2023)), before the pandemic, the Phillips curve correlation was seen as relatively low, suggesting a limited responsiveness of inflation to unemployment. Post pandemic, inflation rates rose well above ECB's 2% target, arguably primarily driven by persistent supply-side shocks, including energy price surges and global supply chain disruptions.

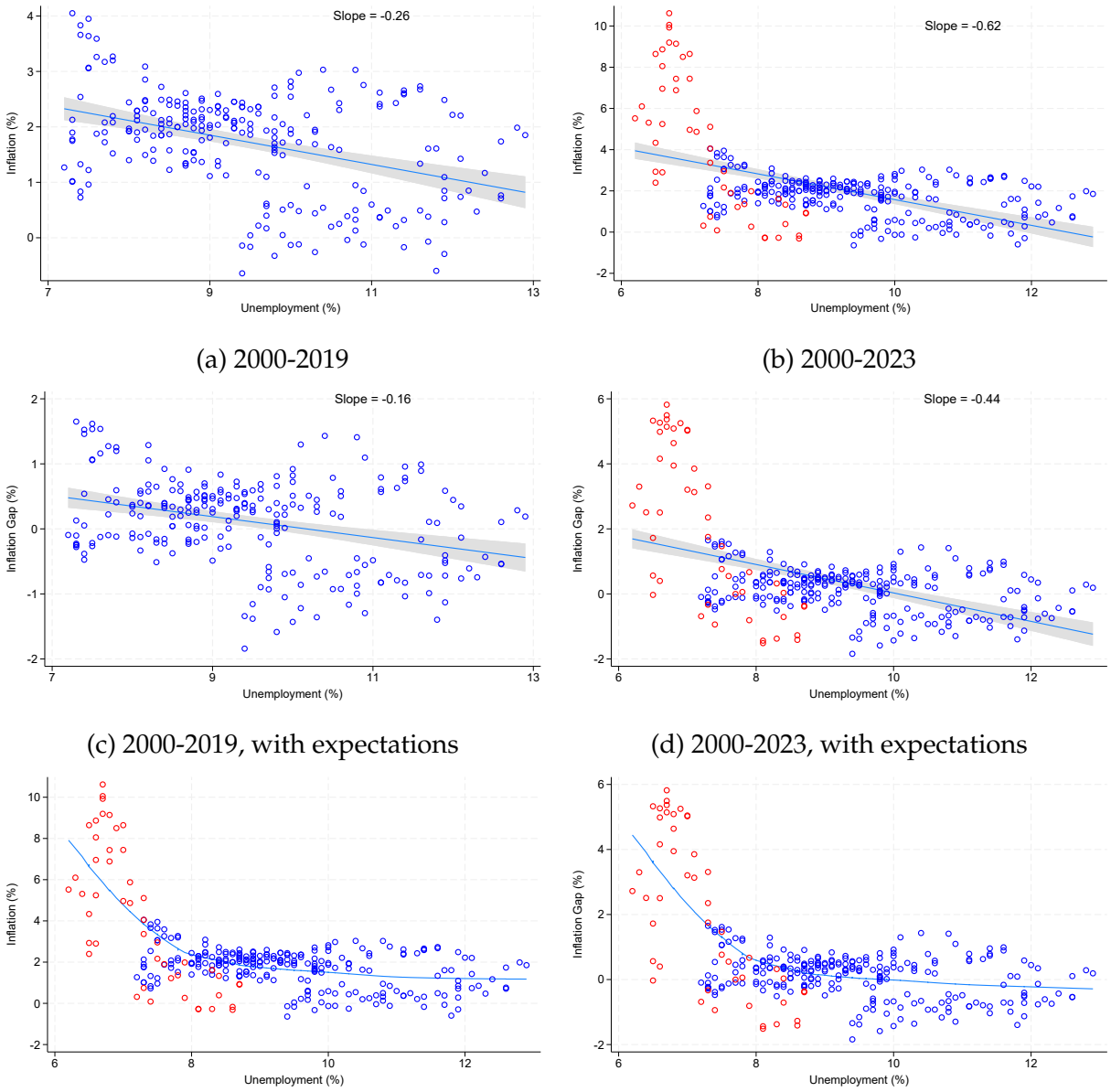
To motivate our econometric analysis, we start by documenting the inflation and unemployment correlation based on various measures and samples in Figure 1. In panels (a) and (b) of Figure 1, we plot a simple Phillips correlation based on monthly aggregate data for EA17, with inflation and unemployment rates based on the pre- and post-COVID19 samples. The 2020-2023 portion of the data, indicated in red in panel (b), show a significant steepening in this period— the slope estimates more than doubles in panel (b). Panel (b) also suggests a sacrifice ratio around 1.25, meaning that reducing the inflation rate by 1 percentage point would require increasing the unemployment rate by approximately 1.25 percentage points.[12]

These patterns remain similar even when accounting for the role of inflation expectations, as we show in panels (c) and (d). In this exercise, we construct an inflation gap measure, defined as the differential between actual inflation (π_t) and expected inflation (π_t^e). For π_t^e, 12-month ahead inflation expectations are based on the ECB's quarterly forecasts, which are then linearly interpolated to a monthly frequency. This inflation gap measure is then plotted against the unemployment rate. Correlations involving this measure also reveal qualitatively similar trade-offs as shown earlier in panels (a) and (b). Quantitatively, the slopes are now slightly flatter.

In panels (e) and (f), we document further the changes in the unemployment and inflation correlation with LOWESS regressions for both inflation and inflation gap measures. This exercise more clearly reveals a recent steepening in the reduced-form Phillips curve relationship.

These correlations are helpful in documenting the recent changes in inflation dynamics in the Euro Area in a reduced-form way, but a formal econometric exercise with an identification strategy is critical to shed light on the accurate magnitude of the inflation-unemployment trade-off and the associated sacrifice ratio. This is especially important to assess whether there indeed has been a true (structural) change in the Phillips curve

[12] We follow the traditional literature where sacrifice ratios have been calculated from an estimated Phillips curve. See Ball (1994).



(a) 2000-2019

(b) 2000-2023

(c) 2000-2019, with expectations

(d) 2000-2023, with expectations

(e) 2000-2023 (LOWESS)

(f) 2000-2023, with expectations (LOWESS)

Figure 1: Phillips Correlation in the Euro Area

Notes: The figure presents the relationship between the inflation rate and unemployment for EA17 aggregates, for 2000m1-2023m12. The slopes in panels (a)-(d) are estimated based on the linear fit of inflation on unemployment in different sample periods. In panels (e)-(f), the slopes are estimated by LOWESS regression. For panels (c), (d), and (f), the inflation gap is defined as $\pi_t - \pi_t^e$, where π_t^e represents 12-month ahead inflation expectations. Observations from 2000-2019 are in blue, and from 2020-2023 are in red. See text in Section-4 for details on variable definitions.

slope in the post COVID period, as this period was subject to persistent supply shocks. A framework that allows controls for the relative price of tradables is needed to account for this issue. We thus next move on to presenting our econometric results based on the panel LP-IV approach.

3 Econometric Results

3.1 Panel Phillips Multiplier Estimation Results

We start by presenting results on the Phillips multiplier, estimated using the panel local projection IV approach.

3.1.1 Baseline Results

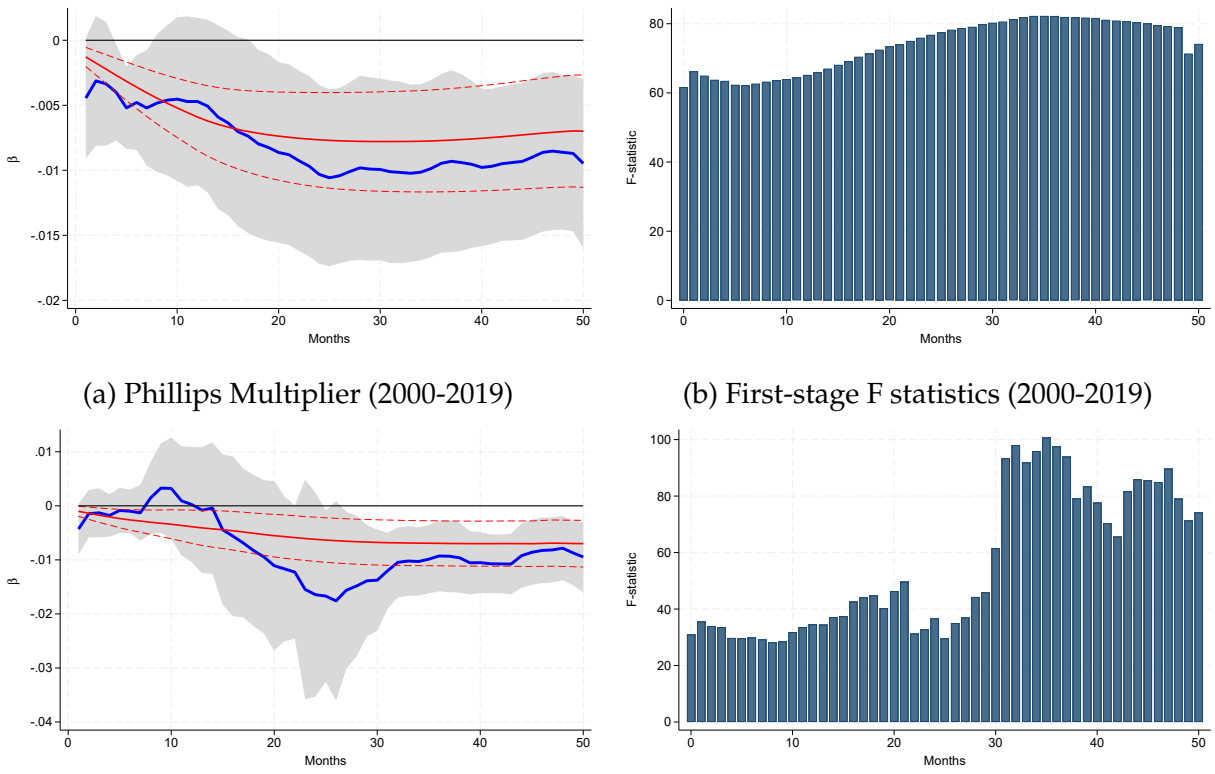
We document highly robust in-sample evidence for the inflation-unemployment trade-off in the Euro Area, by estimating equation (2). We begin by presenting our main results using headline inflation. Figure 2 reports these results for the EA17 countries, covering both the pre- and post-pandemic periods, respectively.

In Figure 2, Panel (a) displays estimates of the Phillips multiplier, β_h, from panel local projections based on equation (2) over horizons $h = 1,\ldots,50$ months. OLS estimates are shown in a solid red line, with their corresponding 90% confidence bands in dashed red lines. The IV estimates are shown in blue, and their 90% confidence bands (the shaded gray area) are based on standard errors clustered at the country level. Panel (b) reports the first-stage results from the panel LP-IV regressions at each horizon, using F-statistics by Montiel Olea and Pflueger (2013).[13]

These results reveal several insights. First, our panel LP-IV approach to estimating the Phillips multiplier in equation (2) yields highly robust results, capturing a strong inflation-unemployment trade-off, particularly beyond a 15- to 20-month horizon.[14] Our IV estimates is around -0.01 at a horizon of 50 months and are statistically significant at the 10% level, based on standard errors clustered at the country level. At horizons where we obtain statistically significant Phillips multiplier estimates, we also observe strong

[13]While Driscoll and Kraay (1998) standard errors are commonly used in settings with cross-sectional correlation, they are not compatible with the Montiel Olea and Pflueger (2013) weak IV robust F-statistics, which we rely on to assess instrument strength with multiple instruments. In Section 3.2, we adopt Driscoll and Kraay (1998) standard errors for forecast evaluation in a panel setting.

[14]In Appendix E, we show how the standard textbook sticky price model augmented with internal sources of dynamics (as developed in Bhattarai et al. (2014)), leads to a dynamic multiplier for demand shocks, analogous to our empirical estimates. Moreover, we show that the Phillips multiplier is monotonically related to the structural Phillips curve slope in the model.



(a) Phillips Multiplier (2000-2019)

(b) First-stage F statistics (2000-2019)

(c) Phillips Multiplier (2000-2023)

(d) First-stage F statistics (2000-2023)

Figure 2: Phillips Multiplier for EA17 Countries (with headline inflation)

Notes: The figure presents results for the EA17 countries over two sample periods: 2000m1-2019m12 (Panels a and b) and 2000m1-2023m12 (Panels c and d). Panel (a) displays the Phillips multiplier estimates, β_h, from panel local projections (equation 2) over horizons $h = 1, \ldots, 50$ months. OLS estimates are shown in red, and IV estimates in blue. The instrument is the U.S. monetary policy shock from Bauer and Swanson (2023), interacted with country fixed effects, denoted $MShock_{n,t}$. Shocks are standardized to have a standard deviation of one. Standard errors are clustered at the country level, and 90% confidence intervals are reported. Panel (b) shows the first-stage F statistics across horizons, based on the robust procedure of Montiel Olea and Pflueger (2013). Panels (c) and (d) replicate the same analysis using the extended sample through 2023.

first-stage results. The Montiel Olea and Pflueger (2013) F statistics exceed 10 (which is a commonly used rule of thumb), and are greater than 20 across almost all horizons and samples, indicating that the instrument is indeed strong.

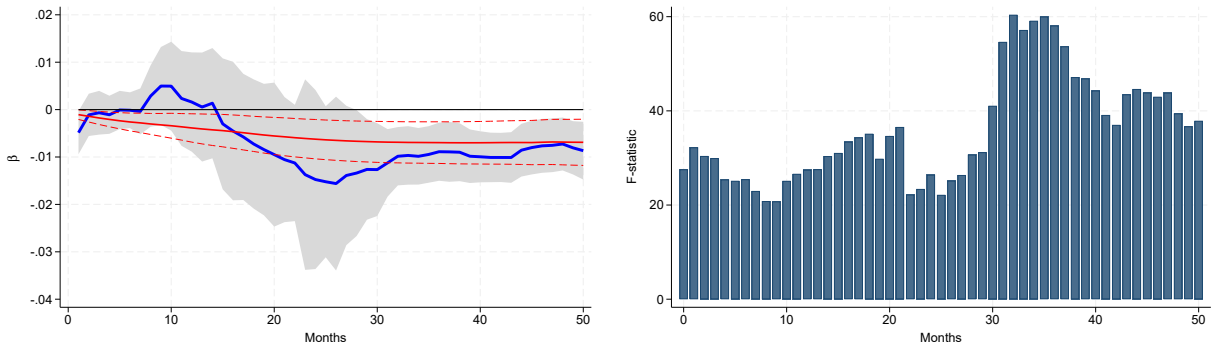
Second, while the IV Phillips multiplier is slightly bigger when including the post-pandemic data, we do not find a substantial difference in Phillips multiplier estimates between the pre- and post-pandemic samples. This contrasts with the much different inflation-unemployment correlations in the two sample periods reported earlier in Figure 1. This also contrasts with the results of Hobijn et al. (2023), who estimate reduced-form Phillips curves at the country level and document a post-pandemic steepening of the relationship between inflation and slack.

Third, our OLS estimates are smaller (in absolute terms) than the IV estimates, consistent with our expectation of a downward bias introduced by supply shocks and endogenous monetary policy responses. The Phillips multiplier estimates with OLS are nevertheless robustly estimated to be around -0.007 at 50 months. Since our strategy leverages data from a currency union, and controls for global supply shocks and common inflation expectations using time fixed effects and country trends, many of the standard identification challenges are partly mitigated even with OLS estimates. Thus, our OLS estimates likely isolate some non-trivial demand-side variation.

3.1.2 Controlling for the relative price of tradables to nontradables

To account for the role of external supply shocks in shaping inflation dynamics, we augment the Phillips Multiplier specification in equation (2) with 12 lags of relative price of tradable to nontradable goods i.e., $p^{TN} = ln(P^T/P^{NT})$. This ratio serves as a proxy for supply-side pressures, particularly those arising from global commodity price shocks, imported inflation, and disruptions in international supply chains. In an open economy, one can postulate that tradable goods prices are heavily influenced by external factors, whereas nontradable prices are more closely tied to domestic labor costs and local demand. An increase in $p^{TN} = ln(P^T/P^{NT})$ typically reflects a rise in tradable prices due to cost-push shocks from abroad. By including lags of this variable as a control in the model, we isolate further externally driven fluctuations in headline inflation, thereby improving identification of the domestic inflation process and enhancing the accuracy of estimated responses to slack and expectations.

This control is especially valuable for the COVID19 period, with large relative price swings, such as energy price surges or global supply bottlenecks. The results for the period that includes post-COVID19 data are therefore presented in Figure 3. As in our benchmark Phillips multiplier model, the augmented framework yields similar estimates



(a) Phillips Multiplier

(b) First-stage F statistics

Figure 3: Phillips Multiplier with Relative Price of Tradables as a Control

Notes: The figure presents results for the EA17 countries over 2000m1-2023m12. The figure displays the Phillips multiplier estimates, β_h, from panel local projections (equation 2) augmented with 12 lags of p^{TN} over horizons $h = 1,\ldots,50$ months. OLS estimates are shown in red, and IV estimates in blue. The instrument is the U.S. monetary policy shock from Bauer and Swanson (2023), interacted with country fixed effects, denoted $MShock_{n,t}$. Shocks are standardized to have a standard deviation of one. Standard errors are clustered at the country level, and 90% confidence intervals are reported.

from the IV approach: at around 25-month horizons, the results become statistically significant and at the 50 month horizon, the IV estimate (blue line) is -0.009. The OLS (red line) estimate is relatively smaller than the IV estimate. We also note that comparing Figure 3 with panel (c) of Figure 2, we find comparable magnitudes of Phillips multiplier estimates at h=50 when including the post-COVID-19 data.

3.1.3 Extensions and sensitivity analysis

Phillips Multiplier with Headline Inflation for the EA12 Panel

In Figure A1 in Appendix, we repeat the analysis focusing on a smaller panel with the initial EA12 members. Our results reveal even more pronounced differences between panel LP-IV and OLS approaches, especially after h=25-30 months. Both in the pre- and post-pandemic samples, the slope estimates are similar. Moreover, for this group of countries, the estimated slopes are slightly larger in magnitude (around -0.015 with panel LP-IV and -0.010 with OLS) compared to their counterparts from the EA17 sample.

Phillips Multiplier with Nontradable (Core) Inflation

One of the identification remedies suggested by McLeay and Tenreyro (2020) is to estimate Phillips curves based on sectoral data. In Figures A2-A3, we report results based on nontradable sector inflation (a measure of core inflation). These estimates can be considered slightly larger compared to the headline inflation results. Also, consistent with the

results under headline inflation: (i) panel LP-IV and OLS strategies yield different slope estimates, (ii) the slope estimates from the EA12 group are larger in magnitude compared to the EA17 group, and (iii) we obtain precise estimates after around h=10-15 months in both of the panel LP methods. Accordingly, at our longest horizon h=50, the EA17 panel yields estimates of around -0.014 (IV) and -0.010 (OLS). With the EA12 panel, the IV and OLS estimates are around -0.021 and -0.015, respectively.

3.1.4 Discussion

Our LP-IV approach yields robust estimates of the inflation-unemployment tradeoff. At a horizon of 50 months, for headline inflation in the EA17 group, and across both the pre- and post-COVID19 periods, we estimate a multiplier of approximately -0.010 within a 90% confidence interval. The corresponding OLS estimate is relatively smaller, at -0.007. For nontradables (core) inflation, the slope appears steeper. The corresponding panel LP-IV estimate is -0.014, also within a 90% confidence band, while the OLS approach produces a coefficient that is relatively smaller.

To put these findings in context, we compare our estimates to those reported by Eser et al. (2020), who study EA inflation using a quarterly panel of pre-pandemic data and focus on core inflation. They employ a dynamic panel approach, with unemployment as their slack measure and run OLS and fixed effects regressions. Eser et al. (2020) obtain slope estimates ranging from -0.005 to -0.014. They also implement a Phillips multiplier model using a LP-IV strategy on monthly *aggregate* Euro Area data, instrumenting with high-frequency monetary policy shocks from Gertler and Karadi (2015). At a 24-month horizon—their longest—they estimate a multiplier of -0.005 with a 68% confidence band. This is around half of the value of our closest comparable estimate of -0.010. The key difference in these results is likely due to our use of a panel LP-IV strategy, as opposed to using aggregate data.

3.2 Out-of-sample Forecasting Results

A puzzle ubiquitous to many countries is that Phillips curve-based models, or more broadly economic models such as DSGE models, have an unstable performance in inflation forecasting. These models outperform naive, univariate models such as the random walk, only occasionally (See Atkeson and Ohanian (2001), Marcellino et al. (2003), Stock and Watson (2008), Faust and Wright (2013), Edge and Gürkaynak (2011), and Bańbura

and Bobeica (2023), among many others.[15]). Recent studies have also suggested that international factors such as global inflation, slack, energy or import prices have information content for inflation in Euro area economies (Béreau et al. (2018) and Moretti et al. (2019)) or in advanced countries (Ciccarelli and Mojon (2010), Kabukçuoğlu and Martínez-García (2018) and Dur and Martínez-García (2020)). Based on the motivation from the literature and the in-sample evidence gained from the panel Phillips multiplier model studied in section 3.1, we turn next to the out-of-sample predictive performance of a panel Phillips curve model.

3.2.1 Models for forecasting inflation using panel data

We consider the following currency union based panel Phillips curve model as our baseline forecasting model:

$$\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}. \tag{7}$$

Here, we follow closely on the baseline monthly panel Phillips curve multiplier model studied in section 2.2. In the forecasting exercises here, we exclude the large set of fixed effects considered in the Phillips multiplier model in order to maintain a relatively parsimonious model structure, consistent with the forecasting literature.[16]

The monthly inflation rate in each month t and country i is calculated as year-over-year changes in the monthly price index, i.e., $\pi_{it} = 100 \times \log(p_{it}/p_{i,t-12})$. This helps us avoid the seasonal adjustment procedure. To forecast h-month ahead inflation in a panel of currency union members, $\pi_{i,t+h}$, we use the unemployment rate u_{it} and past inflation up to 12 lags. We estimate this panel data model by pooled ordinary least squares and obtain common estimates for the Euro Area.

We consider two more panel data models informed by our analysis in section 2.2. Accordingly, we introduce two novel models for inflation forecasting, i.e. a panel AR(12) model and a panel inflation IRF model:

[15] Marcellino et al. (2003) evaluate 4-quarter ahead CPI inflation forecasts for the Euro area and find that aggregation of forecasts from individual country forecasts is better than aggregate forecasts of CPI inflation, without reporting individual country inflation forecasts.

[16] Including fixed effects can lead to overfitting and Nickell bias in small samples and often reduces forecast performance, particularly when the time dimension is limited. For another example of a parsimonious model used in forecasting, see Engel and Wu (2023).

- A panel AR(12) model,

$$\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h} \tag{8}$$

- A panel inflation IRF,

$$\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{N} \beta_n^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h} \tag{9}$$

where $MShock_{i,t}^{n}$ represents instruments for each country $n = 1,...,N$ (with $N = 12$ or $N = 17$) captured by the interaction of Bauer and Swanson (2023) monetary policy shocks with country dummies, defined as in section 2.2.

Finally, we consider

- An AR(12) (univariate time series) model,

$$\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}, \tag{10}$$

as a naive benchmark to evaluate relative performance of models.[17]

Notice that each of these panel models (i.e. models (7), (8), or (9)) is non-nested with the univariate AR(12) model (10).

3.2.2 Forecast scheme, sample, and evaluation

We adopt a rolling window scheme, where the first estimation window spans the 2000m1-2007m12 period (hence each window covers 96 months). Our baseline forecast horizon is 12 months.[18] The forecasting period starts in 2008m12 and ends in 2019m12, i.e. before the COVID-19 pandemic era. We also consider an alternative sample ending in 2023m12. The forecasting sample covers the Euro Area debt crisis starting from 2009 to late 2010s. In terms of the country groups, we focus on EA17 samples, constructed based on the 17 members of EA. We adopt headline inflation as our baseline measure, while we late separately also report results from nontradable sector inflation.

[17] Our initial assessment suggests that the AR(12) model outperforms the 12-month random walk or 12-month moving moving average in terms of the RMSE metric.

[18] We also report 18 and 24 month ahead forecasts in extensions.

Our forecast evaluation strategy is as follows. We construct the relative root mean square error (RMSE) metric of the baseline PC panel model 7 (or other panel models (8), or (9)) against the naive univariate benchmark. We test the null hypothesis of equal predictive accuracy, against the alternative where a panel model is more accurate than the naive model. We then perform Diebold and Mariano (1995) and West (1996) (DMW) test for the non-nested models. We report Driscoll and Kraay (1998) standard errors robust to cross-sectional and temporal correlations, especially as the time dimension in the panel is relatively large. We allow for 12 lags in the serial correlation of error terms.

3.2.3 Forecasting results

In Table 1, we present forecasting results for headline inflation for EA17 countries in the pre-2019 sample. The first column reports the absolute RMSEs from the AR(12) process, which are notably low across all countries. Critically, the panel models generally outperform this univariate benchmark for the majority of EA17 members, as shown in panels (a)-(c) through the relative RMSE metric. These improvements are statistically significant according to the DMW test results in most cases—and comparable across all three panel models—highlighting the advantages of pooling information across countries in a currency union to enhance inflation forecasts at the national level.

Table 2 extends the analysis to the post-pandemic period, covering data through 2023:m12. Forecast errors from the naive AR(12) benchmark across all countries increase substantially during this period, reflecting heightened uncertainty. However, the panel models' good forecasting performance remains robust, with them continuing to beat the AR(12) benchmark both in terms of relative RMSE and DMW test statistics.[19]

To illustrate the forecast performance of our novel panel approach, Figures 4-6 compare the Panel PC model's one-year ahead forecasts with actual, monthly, year-over-year inflation rates for all the EA17 countries (i.e., we compare the forecasts made 12 months prior to time t with the actual inflation observed at time t.) Overall, the panel Phillips curve model tracks inflation rates remarkably well over the full sample period, including the low-inflation era from 2008 to the onset of the pandemic—a period of particular interest due to the missing inflation and disinflation puzzles (see Ball and Mazumder (2021)).[20] An exception arises during the COVID-19 episode, where the model slightly

[19] According to the Schwarz Information Criterion (SIC), the panel AR(1) specification yields the lowest SIC value among panel AR(p) models for p=1,...,12. However, in terms of forecast performance measured by RMSE, the panel AR(12) provides the best results.

[20] Inflation declined less than predicted by the textbook Phillips curve during Europe's 2008 and 2011 recessions, and has risen less than expected during subsequent recoveries.

Table 1: Out-of-sample forecast performance of panel models with headline inflation for 2008m12:2019m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{17} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

		(a)			(b)			(c)		
Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value	Rel. RMSE Panel AR	DMW stat.	DMW p-value	Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.03	**0.91***	1.59	0.06	**0.91***	1.62	0.05	**0.91***	1.42	0.08
Belgium	0.04	**0.91****	1.72	0.04	**0.90****	1.73	0.04	**0.90****	1.76	0.04
Cyprus	0.06	**0.84****	1.92	0.03	**0.85****	1.88	0.03	**0.85****	1.86	0.03
Estonia	0.05	**0.91*****	2.46	0.01	**0.91*****	2.41	0.01	**0.91*****	2.68	0.00
Finland	0.03	**0.90****	2.31	0.01	**0.90*****	2.35	0.01	**0.90****	2.32	0.01
France	0.03	**0.69***	1.31	0.10	**0.68***	1.31	0.10	**0.68***	1.32	0.09
Germany	0.03	1.09	-1.97	0.98	1.08	-1.66	0.95	1.08	-1.67	0.95
Greece	0.06	**0.77****	1.92	0.03	**0.79****	1.79	0.04	**0.79****	1.76	0.04
Ireland	0.03	**0.94**	0.94	0.18	**0.93**	1.07	0.14	**0.94**	0.98	0.16
Italy	0.03	**0.80****	2.30	0.01	**0.81****	2.25	0.01	**0.81****	2.19	0.01
Luxembourg	0.05	**0.86****	2.27	0.01	**0.85****	2.28	0.01	**0.85*****	2.33	0.01
Malta	0.05	**0.93**	0.87	0.19	**0.92**	0.89	0.19	**0.93**	0.85	0.20
Netherlands	0.03	**0.94***	1.52	0.06	**0.94***	1.57	0.06	**0.94***	1.49	0.07
Portugal	0.04	1.01	-0.24	0.60	1.01	-0.34	0.63	1.02	-0.43	0.67
Slovakia	0.03	**0.86*****	4.72	0.00	**0.86*****	4.62	0.00	**0.86*****	4.40	0.00
Slovenia	0.04	**0.87****	2.19	0.01	**0.86****	2.21	0.01	**0.86****	2.21	0.01
Spain	0.06	**0.63***	1.48	0.07	**0.64***	1.46	0.07	**0.64***	1.46	0.07

Notes: This table reports results for headline inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2019m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 2: Out-of-sample forecast performance of panel models with headline inflation for 2008m12:2023m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{17} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

		(a)			(b)			(c)		
Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value	Rel. RMSE Panel AR	DMW stat.	DMW p-value	Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.05	**0.88****	1.66	0.05	**0.86****	1.74	0.04	**0.85****	1.69	0.05
Belgium	0.07	**0.82****	1.70	0.04	**0.81****	1.72	0.04	**0.81****	1.72	0.04
Cyprus	0.07	**0.85****	2.16	0.02	**0.84****	2.16	0.02	**0.84****	2.14	0.02
Estonia	0.08	1.03	-0.27	0.61	**0.99**	0.11	0.46	**0.99**	0.15	0.44
Finland	0.04	**0.90****	2.31	0.01	**0.88*****	2.55	0.00	**0.88*****	2.47	0.01
France	0.04	**0.83***	1.28	0.10	**0.81***	1.42	0.08	**0.82***	1.37	0.09
Germany	0.05	**0.88**	1.11	0.13	**0.87**	1.13	0.13	**0.87**	1.10	0.13
Greece	0.07	**0.85*****	2.34	0.01	**0.84*****	2.38	0.01	**0.85****	2.25	0.01
Ireland	0.06	**0.73****	1.65	0.05	**0.72****	1.68	0.05	**0.72***	1.64	0.05
Italy	0.06	**0.84*****	2.39	0.01	**0.84****	2.21	0.01	**0.84****	2.19	0.01
Luxembourg	0.07	**0.86****	2.20	0.01	**0.85*****	2.34	0.01	**0.84****	2.29	0.01
Malta	0.05	**0.82****	1.80	0.04	**0.82****	1.81	0.04	**0.83****	1.76	0.04
Netherlands	0.08	0.96	0.52	0.30	0.96	0.56	0.29	0.96	0.54	0.30
Portugal	0.06	**0.81***	1.57	0.06	**0.81***	1.54	0.06	**0.81***	1.52	0.07
Slovakia	0.04	1.12	-0.83	0.80	1.07	-0.60	0.73	1.06	-0.56	0.71
Slovenia	0.05	**0.98**	0.40	0.35	**0.97**	0.69	0.24	**0.98**	0.52	0.30
Spain	0.07	**0.78***	1.64	0.05	**0.76****	1.79	0.04	**0.76****	1.81	0.04

Notes: This table reports results for headline inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2023m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

underpredicts inflation, reflecting the extraordinary nature of the shock and the heightened uncertainty in that period.

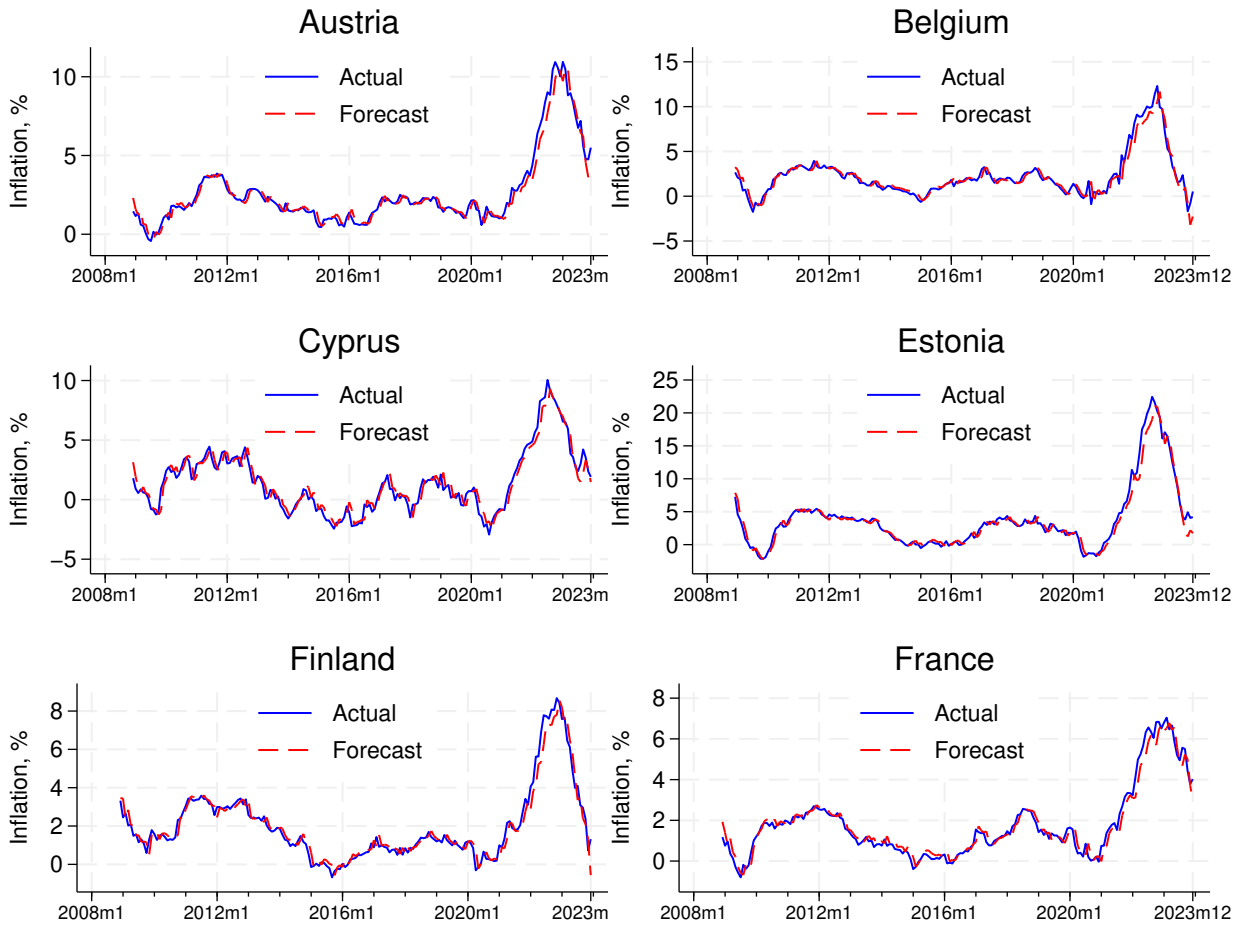


Figure 4: Panel Phillips curve forecasts vs. actual headline inflation 2008m:12-2023:m12

Notes: This figure plots headline inflation forecasts vs. actual inflation for 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

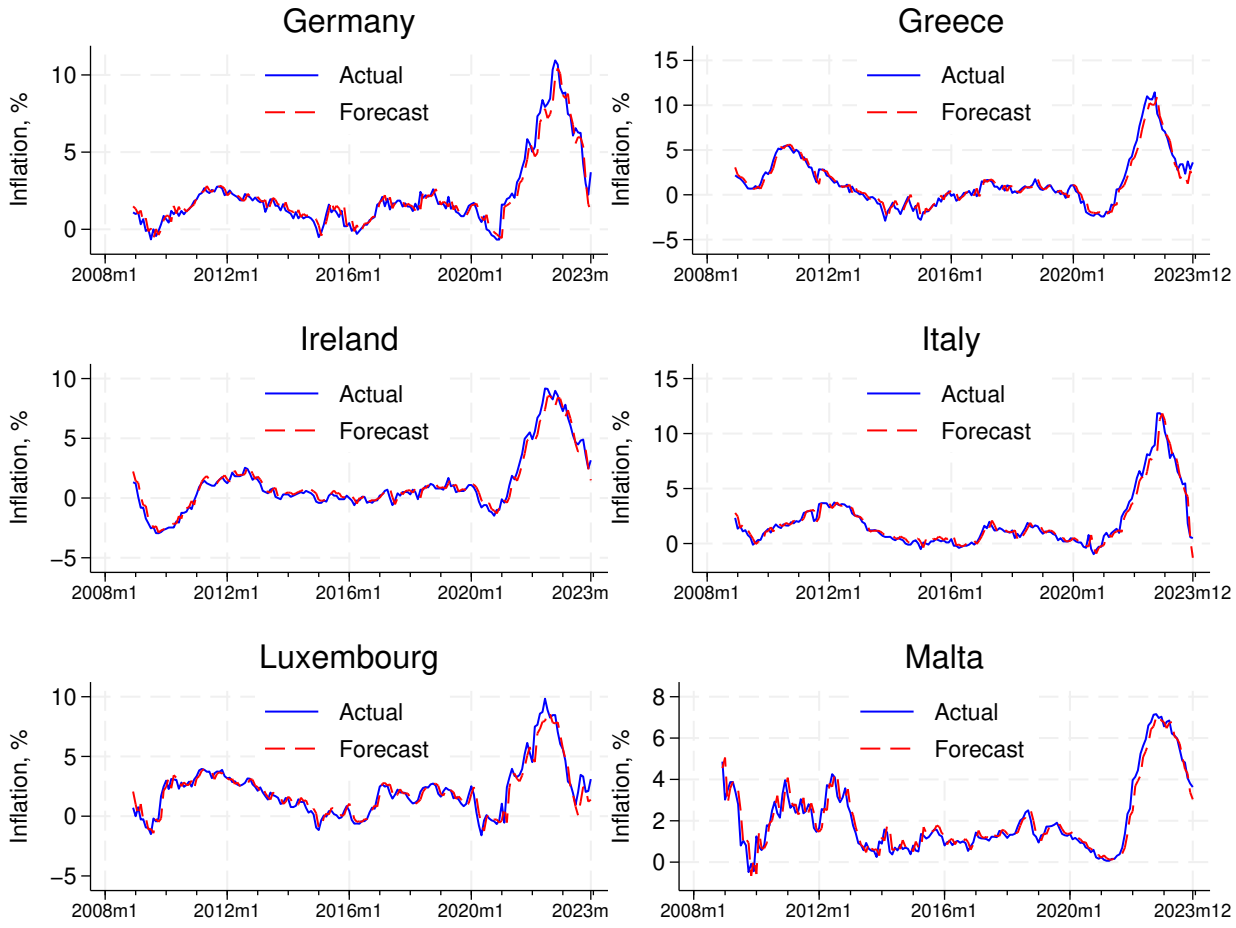


Figure 5: Panel Phillips curve forecasts vs. actual headline inflation 2008m:12-2023:m12

Notes: This figure plots headline inflation forecasts vs. actual inflation for 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

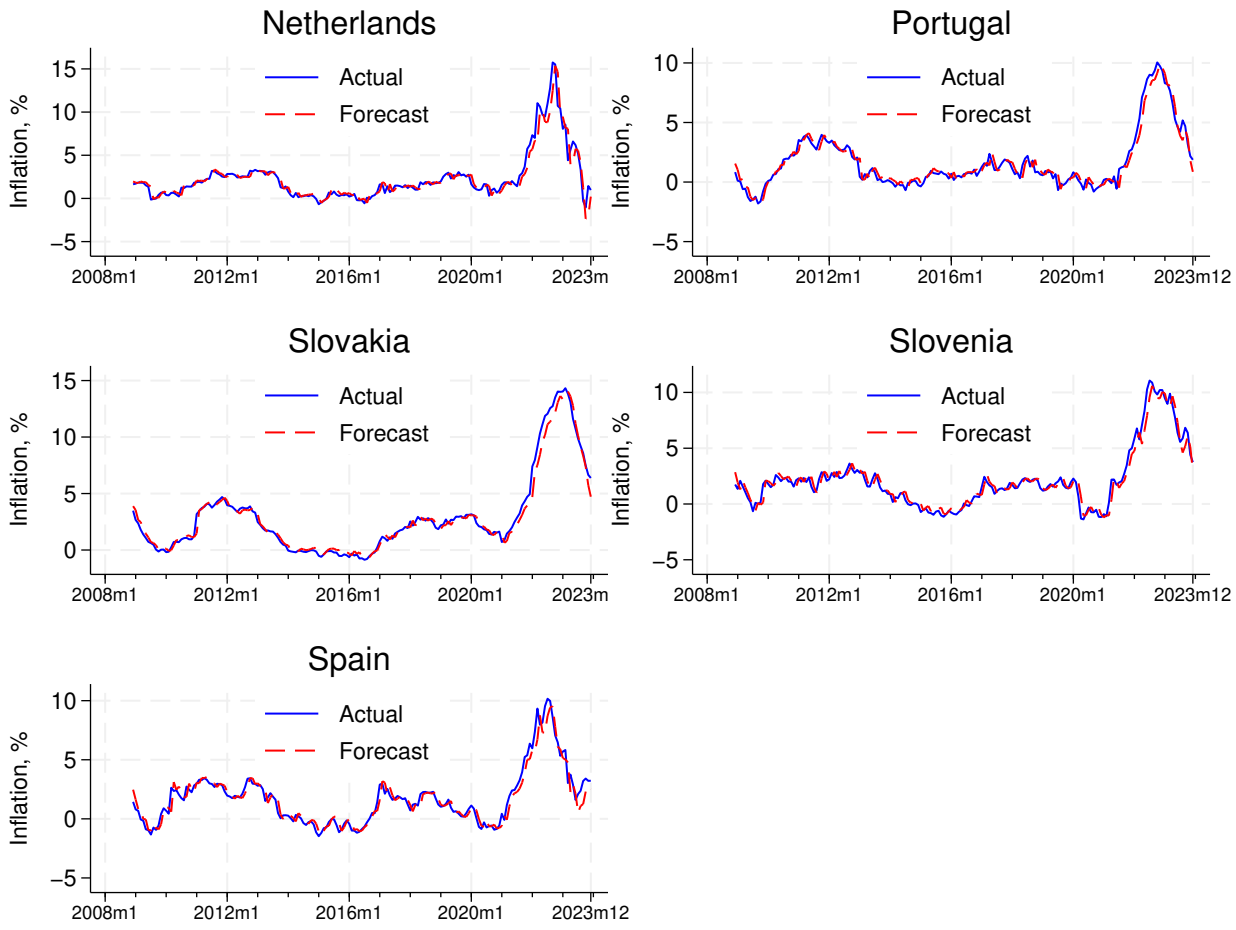


Figure 6: Panel Phillips curve forecasts vs. actual headline inflation 2008m:12-2023:m12

Notes: This figure plots headline inflation forecasts vs. actual inflation for 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

3.2.4 Controlling for the relative price of tradables to nontradables

In line with section 3.1 where we considered an extension to the Phillips Multiplier model by incorporating the relative price of tradable to nontradable goods, we consider the following panel Phillips curve (PC-2) model for inflation forecasting:

$$\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \sum_{j=1}^{12} \beta_j p^{TN}_{i,t-j} + \varepsilon_{i,t+h} \tag{11}$$

where we include 12 lags of relative prices of tradable and nontradable sectors, denoted $p^{TN} = log(P^T / P^{NT})$. The forecast results from this panel model are presented in Table 3 and they are comparable to the baseline PC model.

3.2.5 Extensions and Sensitivity Analysis

Nontradable (core) inflation panel forecasts: Next, we turn to forecasts for the nontradables sector in EA17, presented in Tables A1 and A2 for the pre- and post-2019 periods, respectively. These results demonstrate that panel models retain strong predictive power in this sector, particularly over the full sample including the post-pandemic period, and consistently outperform the AR(12) benchmark as measured by the relative RMSE. We also plot our forecasts vis-a-vis the actual inflation rates for the nontradables sector in Figures A4-A6. Again, the key results and the main message remain robust.

Alternative forecast horizons in headline inflation panel forecasts: Longer forecast horizons are particularly relevant for the ECB's medium-term projections and broader policy objectives. Tables A3 and A4 report results for extended forecast horizons of h=18 and h=24 months, respectively. As the results indicate, forecast accuracy remains high for the majority of EA17 countries even at these extended horizons.

Headline inflation EA12 panel forecasts: In Tables A5 and A6, we restrict the analysis to EA12 countries for both the pre- and post-pandemic periods. The key takeaway is that while panel models continue to outperform the univariate AR(12) benchmark, the forecasting gains are more pronounced when EA17 countries are pooled together in a panel data framework compared to EA12, suggesting that a broader information base has greater predictive content.

Forecasts based on individual-country Phillips curves: We also explore an alternative specification based on individual-country Phillips curve models, in the spirit of Atkeson and Ohanian (2001). To ensure comparability with our panel PC specification, each coun-

Table 3: Out-of-sample forecast performance of panel models with headline inflation for 2008m12:2023m12 (h=12)
(a) Phillips curve with relative price of tradables and nontradables (PC-2):
$\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \sum_{j=1}^{12} \beta_j p_{i,t-j}^{TN} + \varepsilon_{i,t+h}$
(b) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

	(a) Headline 2019				(b) Headline 2023			
Country	AR(12) RMSE	RMSE Panel PC-2	DMW stat.	DMW p-value	AR(12) RMSE	RMSE Panel PC-2	DMW stat.	DMW p-value
Austria	0.03	**0.91***	1.62	0.05	0.05	**0.86****	1.66	0.05
Belgium	0.04	**0.90****	1.80	0.04	0.07	**0.83****	1.93	0.03
Cyprus	0.06	**0.84****	1.87	0.03	0.07	**0.85****	1.98	0.02
Estonia	0.05	**0.90*****	2.60	0.01	0.08	1.02	-0.19	0.58
Finland	0.03	**0.90****	1.93	0.03	0.04	**0.88****	2.25	0.01
France	0.03	**0.70**	1.24	0.11	0.04	**0.83**	1.26	0.10
Germany	0.03	1.10	-2.05	0.98	0.05	**0.87**	1.06	0.14
Greece	0.06	**0.78****	1.94	0.03	0.07	**0.85*****	2.41	0.01
Ireland	0.03	**0.95**	0.81	0.21	0.06	**0.72***	1.60	0.06
Italy	0.03	**0.81****	2.17	0.02	0.06	**0.83****	2.27	0.01
Luxembourg	0.05	**0.85*****	2.35	0.01	0.07	**0.86****	2.13	0.02
Malta	0.05	**0.95**	0.55	0.29	0.05	**0.86***	1.49	0.07
Netherlands	0.03	**0.94***	1.63	0.05	0.08	**0.94**	0.59	0.28
Portugal	0.04	1.01	-0.18	0.57	0.06	**0.82***	1.56	0.06
Slovakia	0.03	**0.85*****	4.83	0.00	0.04	1.15	-0.88	0.81
Slovenia	0.04	**0.86****	2.24	0.01	0.05	**0.97**	0.66	0.26
Spain	0.06	**0.63***	1.50	0.07	0.07	**0.77****	1.69	0.05

Notes: This table reports results for headline inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2023m12 in panel (a) and in 2019m12 in panel (b). We report the relative RMSE of the Panel Phillips Curve (PC-2) model augmented with the relative price of tradables and nontradables with 12 lags vs. the AR(12) benchmark. The first column in each panel (a) and (b) presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

try's model includes the same set of regressors: current-period unemployment and 12 lags of inflation. We assess forecast performance relative to the AR(12) benchmark across both headline and nontradables inflation series, and for both pre- and post-COVID periods. Since the Phillips curve model nests the AR(12), we rely on the Clark and West (2007) (CW) test, which accounts for the potential bias arising from the noise introduced by additional regressors in the larger model. Despite this adjustment, the Phillips curve model performs generally worse than the AR(12) benchmark, suggesting that the Atkeson and Ohanian (2001) "Phillips curve puzzle" remains pervasive across the EA17 sample when using individual country Phillips curves. See Table A7 for results based on both headline and core inflation Phillips curves.

Forecast evaluation in unstable environments: The standard Diebold and Mariano (1995) and West (1996) tests evaluate forecast performance over the entire sample which has been subject to instabilities due to local and global economic factors. We consider the possibility of unstable environments in which the relative forecast accuracy of two models may vary across subsamples. A prominent approach for formally evaluating forecasts under such instabilities is the fluctuation test proposed by Giacomini and Rossi (2010).[21] Here, we apply this test to headline inflation forecasts in our baseline exercise using 12-month-ahead forecasts from (i) an AR(12) time-series model and (ii) the panel Phillips curve model.

Unlike the previous sections, we adopt a different forecast evaluation metric. Following Giacomini and Rossi (2010), we compute local, out-of-sample mean squared forecast error (MSFE) differences between the AR(12) model and the panel Phillips curve in 10-year rolling windows, beginning with the first point forecast in 2008:m12. This approach allows us to assess the models' *local* relative performances during various subsamples encompassing episodes of low and high inflation in the Euro Area, both before and after the COVID-19 pandemic.

In (solid) blue line, Figures A7-A9 display the (standardized) local relative MSFE of the AR(12) time-series model and the panel PC model. The (dashed) red line in each figure represents the critical value for testing the null hypothesis that the two models have equal out-of-sample performance in each subsample, against the alternative that the panel Phillips curve performs better in at least one subsample. When the local relative MSFE exceeds the critical value, we reject the null hypothesis and conclude that the panel Phillips curve delivered more accurate forecasts than the AR(12) model during those periods.

Figures A7-A9 report fluctuation test statistics that are generally above zero, indicat-

[21] See Rossi (2021) for a review of forecast techniques in the presence of instabilities.

ing that the panel PC outperforms the AR(12) model in most Euro Area (EA17) countries. In assessing the relative forecast performance in recent years, including the pandemic period, the text yields mixed results across countries. In countries such as Germany and Ireland, the panel Phillips curve achieved greater accuracy than the AR(12) model post-pandemic. In others, such as Austria and Italy, the test statistic declined after the pandemic, but remained positive—indicating that the panel Phillips curve still outperformed the AR(12) model even in the post-pandemic period.

The critical values (in dashed red lines) indicate statistical significance at the 10% level. Accordingly, the fluctuation test identifies at least one subsample in which the panel PC model performs significantly better than the AR(12) model in the majority of EA17 countries (with France, Germany, Ireland, the Netherlands, and Portugal as exceptions). The COVID-19 period did not cause a systematic deterioration in the forecast accuracy of the panel Phillips curve model relative to the AR(12) benchmark. In some countries, performance improved; in others, it declined. Overall, no consistent cross-country pattern emerges. Despite these fluctuations, the panel PC model continues to demonstrate strong and generally superior forecasting performance.[22]

Forecasts with an alternative seasonal adjustment filter: In our baseline analysis, we used the unemployment rate series, UNE-RT-M (SA) from Eurostat, to construct a fully balanced panel for EA17. One concern, particularly in forecasting applications, is that the two-sided filter used for seasonal adjustment in this series may bias the forecast results, as it incorporates information from future observations into the current dataset.[23] To address this, we implement a one-sided moving average filter on the raw unemployment series, UNE-RT-M (NSA), subject to data availability constraints. This one-sided filtering method results in the loss of 11 initial observations from each series available. With this limitation, we construct a panel that includes the broadest possible set of countries and observations. All unemployment series begin in 2000:m1, consistent with the baseline, with the following exceptions: Cyprus starts in 2000:m12, Estonia in 2001:m1, France in 2003:m12, and Germany in 2007:m12. Because the estimation period in our analysis also ends in 2007:m12, we exclude Germany from this robustness exercise.

Our analysis reveals that our results are still robust under the one-sided seasonal ad-

[22] Bańbura et al. (2024) compare inflation forecast performance for the U.S. and the Euro Area and find no robust evidence that one region is consistently easier or more difficult to forecast. The relative forecast performance depends on the sample period and the presence of instabilities—particularly whether it includes episodes like the Global Financial Crisis—and the specific models employed. A key conclusion is that, for both regions, univariate benchmark models are still difficult to outperform.

[23] Eurostat applies a number of seasonal adjustment techniques, including TRAMO-SEATS and X-12-ARIMA. For details in Eurostat's methodology, see: https://ec.europa.eu/eurostat/documents/3859598/5910549/KS-RA-09-006-EN.PDF

justment filter. We present the forecast results with headline inflation in Table A8 and Figures (A10)-(A12). Our relative RMSEs and DMW test results are virtually the same as the baseline exercise. We conclude that data limitations do not bias our baseline forecast results and there is a strong information content for Euro Area inflation from panel data.

4 Conclusion

This paper revisits the inflation-unemployment trade-off and provides novel empirical models to study the inflation dynamics in the Euro Area. We begin by addressing long-standing identification challenges associated with the Phillips curve. Using a panel local projection framework where we model a link between cumulative inflation and unemployment in the spirit of Barnichon and Mesters (2021), we estimate a dynamic Phillips multiplier for 17 Euro Area countries. We leverage U.S. monetary policy shocks by Bauer and Swanson (2023) to construct an IV for unemployment. Our results provide robust evidence of a statistically and economically significant trade-off between inflation and unemployment, particularly at horizons beyond 15 months. The estimates are remarkably consistent across pre- and post-pandemic periods and between core and headline inflation. Furthermore, we find that OLS and IV estimates align closely—suggesting that the use of panel LP framework for a currency union mitigates many of the estimation challenges traditionally associated with the Phillips curve.

In addition to identification, our study contributes to the literature on inflation forecasting, where panel models—to our knowledge—has never been considered before. We show that a panel Phillips curve model offers considerable predictive power at 12-month horizons relative to the univariate AR(12) process, for 17 EA member states and two distinct inflation measures. Although forecasting uncertainty increased slightly post-pandemic, the panel Phillips curve still performs well relative to the naive alternatives. We obtain alternative panel specifications based on the underlying mechanism for the Phillips multiplier, which contain significant information content for inflation.

References

Afrouzi, H., Bhattarai, S., and Wu, E. (2024). Relative-price changes as aggregate supply shocks revisited: Theory and evidence. *Journal of Monetary Economics*, 148(S).

Atkeson, A. and Ohanian, L. (2001). Are Phillips Curves Useful for Forecasting Inflation? *Federal Reserve bank of Minneapolis quarterly review*, 25(1):2–11.

Auerbach, A. J. and Gorodnichenko, Y. (2012). Fiscal Multipliers in Recession and Expansion. In *Fiscal Policy after the Financial Crisis*, pages 63–98. University of Chicago Press.

Ball, L. (1994). What Determines the Sacrifice Ratio? In *Monetary policy*, pages 155–193. The University of Chicago Press.

Ball, L. and Mazumder, S. (2021). A Phillips Curve for the Euro Area. *International Finance*, 24(1):2–17.

Bańbura, M. and Bobeica, E. (2023). Does the Phillips Curve Help to Forecast Euro Area Inflation? *International Journal of Forecasting*, 39(1):364–390.

Bańbura, M., Lenza, M., and Paredes, J. (2024). Forecasting Inflation in the US and in the Euro Area. *Handbook of Research Methods and Applications in Macroeconomic Forecasting*, pages 218–239.

Barnichon, R. and Mesters, G. (2021). The Phillips Multiplier. *Journal of Monetary Economics*, 117:689–705.

Barnichon, R. and Shapiro, A. H. (2024). Phillips Meets Beveridge. *Journal of Monetary Economics*, 148:103660.

Bauer, M. D. and Swanson, E. T. (2023). A Reassessment of Monetary Policy Surprises and High-Frequency Identification. *NBER Macroeconomics Annual*, 37(1):87–155.

Benigno, P., Canofari, P., Di Bartolomeo, G., and Messori, M. (2023). *Inflation Dynamics and Monetary Policy in the Euro Area*. Luiss School of European Political Economy.

Béreau, S., Faubert, V., and Schmidt, K. (2018). Explaining and Forecasting Euro Area Inflation: the Role of Domestic and Global Factors.

Bhattarai, S., Lee, J. W., and Park, W. Y. (2014). Price indexation, habit formation, and the Generalized Taylor Principle. *Journal of Economic Dynamics and Control*, 48(C):218–225.

Ciccarelli, M. and Mojon, B. (2010). Global inflation. *The Review of Economics and Statistics*, 92(3):524–535.

Clark, T. E. and West, K. D. (2007). Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. *Journal of Econometrics*, 138(1):291–311.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 20(1):134–144.

Driscoll, J. C. and Kraay, A. C. (1998). Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. *Review of Economics and Statistics*, 80(4):549–560.

Dur, A. and Martínez-García, E. (2020). Mind the gap!—A Monetarist View of the Open-Economy Phillips Curve. *Journal of Economic Dynamics and Control*, 117:103959.

Edge, R. M. and Gürkaynak, R. S. (2011). How Useful are Estimated DSGE Model Forecasts?

Engel, C. and Wu, S. P. Y. (2023). Liquidity and Exchange Rates: An Empirical Investigation. *The Review of Economic Studies*, 90(5):2395–2438.

Eser, F., Karadi, P., Lane, P. R., Moretti, L., and Osbat, C. (2020). The Phillips Curve at the ECB. *The Manchester School*, 88:50–85.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. In *Handbook of economic forecasting*, volume 2, pages 2–56. Elsevier.

Fitzgerald, T., Jones, C., Kulish, M., and Nicolini, J. P. (2024). Is There a Stable Relationship between Unemployment and Future Inflation? *American Economic Journal: Macroeconomics*, 16(4):114–142.

Forbes, K. J. (2019). Inflation Dynamics: Dead, Dormant, or Determined Abroad? *Brookings Papers on Economic Activity*, (2 (Fall)):257–338.

Furlanetto, F. and Lepetit, A. (2024). The Slope of the Phillips Curve. Technical report, Board of Governors of the Federal Reserve System (US).

Gabriel, R. D. (2023). Monetary Policy and the Wage Inflation-Unemployment Tradeoff. *European Economic Review*, 159:104565.

Gertler, M. and Karadi, P. (2015). Monetary Policy Surprises, Credit Costs, and Economic Activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Giacomini, R. and Rossi, B. (2010). Forecast Comparisons in Unstable Environments. *Journal of Applied Econometrics*, 25(4):595–620.

Hazell, J., Herreno, J., Nakamura, E., and Steinsson, J. (2022). The Slope of the Phillips Curve: Evidence from U.S. States. *The Quarterly Journal of Economics*, 137(3):1299–1344.

Hobijn, B., Miles, R., Royal, J., and Zhang, J. (2023). The Recent Steepening of Phillips Curves. *Chicago Fed Letter*, 475.

Inoue, A., Rossi, B., and Wang, Y. (2024). Has the Phillips Curve Flattened? Technical report.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American economic review*, 95(1):161–182.

Jordà, Ò. and Nechio, F. (2018). Inflation globally.

Jordà, Ò. and Taylor, A. M. (2025). Local Projections. *Journal of Economic Literature*, 63(1):59–110.

Kabukçuoğlu, A. and Martínez-García, E. (2018). Inflation as a Global Phenomenon—Some Implications for Inflation Modeling and Forecasting. *Journal of Economic Dynamics and Control*, 87:46–73.

Marcellino, M., Stock, J. H., and Watson, M. W. (2003). Macroeconomic Forecasting in the Euro Area: Country Specific versus Area-Wide Information. *European Economic Review*, 47(1):1–18.

Mavroeidis, S., Plagborg-Møller, M., and Stock, J. H. (2014). Empirical Evidence on Inflation Expectations in the New Keynesian Phillips Curve. *Journal of Economic Literature*, 52(1):124–188.

McLeay, M. and Tenreyro, S. (2020). Optimal Inflation and the Identification of the Phillips Curve. *NBER Macroeconomics Annual*, 34(1):199–255.

Montiel Olea, J. L. and Pflueger, C. (2013). A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics*, 31(3):358–369.

Moretti, L., Onorante, L., and Saber, S. Z. (2019). *Phillips Curves in the Euro Area*. Number 2295. ECB Working Paper.

Nakamura, E. and Steinsson, J. (2014). Fiscal Stimulus in a Monetary Union: Evidence from US Regions. *American Economic Review*, 104(3):753–792.

Ramey, V. A. and Zubairy, S. (2018). Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. *Journal of political economy*, 126(2):850–901.

Rossi, B. (2021). Forecasting in the Presence of Instabilities: How we Know Whether Models Predict Well and How to Improve Them. *Journal of Economic Literature*, 59(4):1135–1190.

Smith, S. C., Timmermann, A., and Wright, J. H. (2025). Breaks in the Phillips Curve: Evidence from Panel Data. *Journal of Applied Econometrics*, 40(2):131–148.

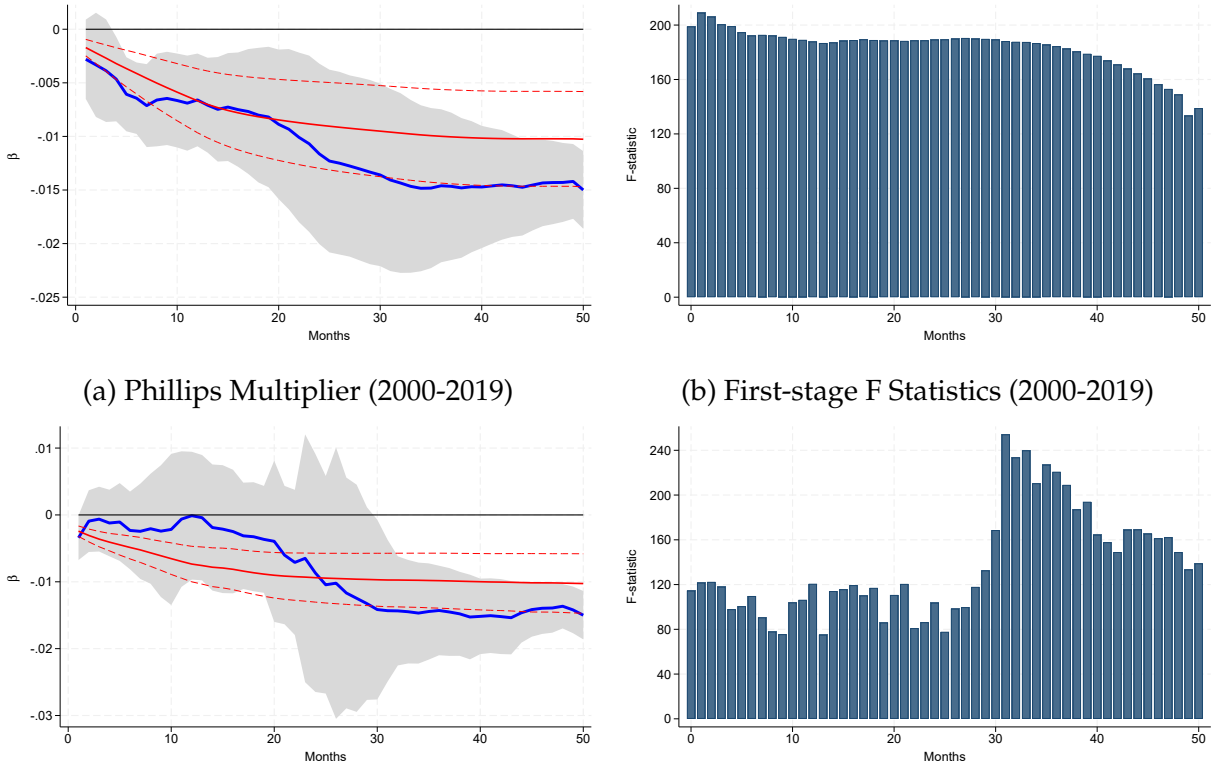
Stock, J. H. and Watson, M. W. (2008). Phillips Curve Inflation Forecasts.

Wellmann, S. (2023). The Phillips Curve in the Euro Area: New Evidence using Country-Level Data.

West, K. D. (1996). Asymptotic Inference about Predictive Ability. *Econometrica: Journal of the Econometric Society*, pages 1067–1084.

Appendix

A. Phillips Multiplier with Headline Inflation for the EA12 Panel



(a) Phillips Multiplier (2000-2019)

(b) First-stage F Statistics (2000-2019)

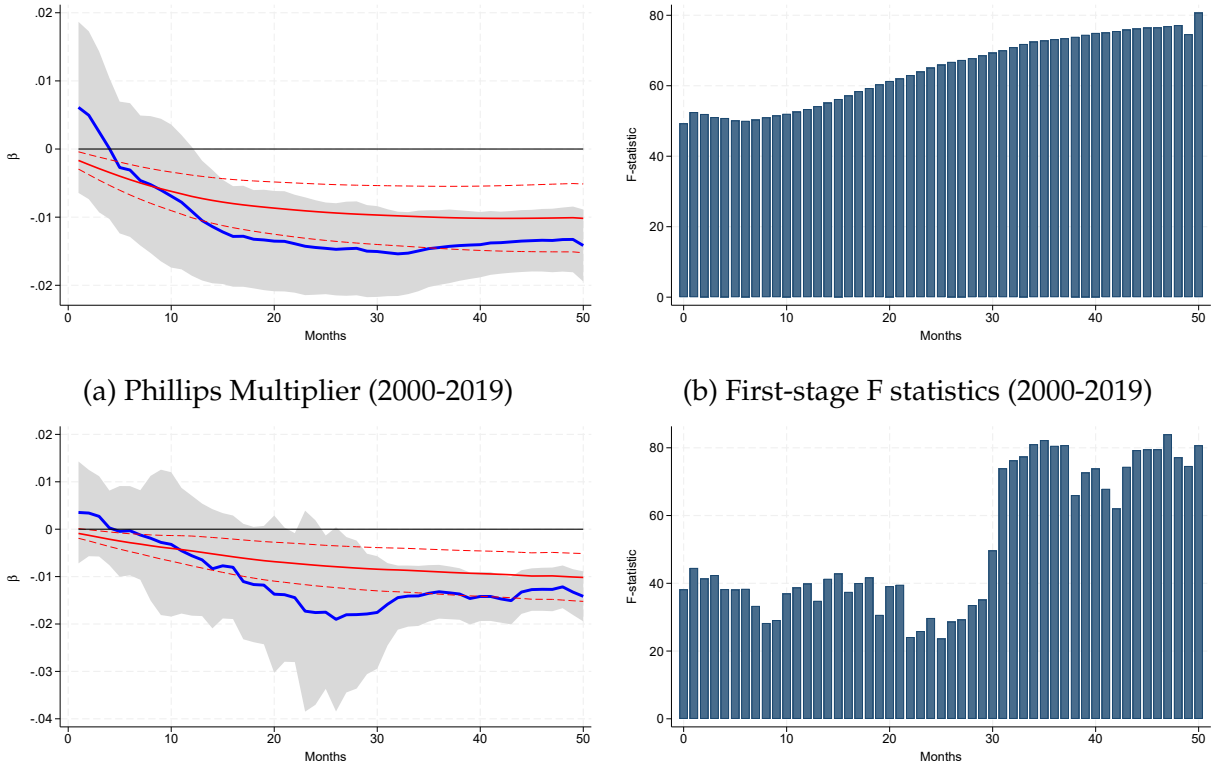
(c) Phillips Multiplier (2000-2023)

(d) First-stage F Statistics (2000-2023)

Figure A1: Phillips Multiplier for EA12 Countries with headline inflation

Notes: The figure presents results for the EA12 countries over two sample periods: 2000m1-2019m12 (Panels a and b) and 2000m1-2023m12 (Panels c and d). Panel (a) displays the Phillips multiplier estimates, β_h, from panel local projections (equation 2) over horizons $h = 1,\ldots,50$ months. OLS estimates are shown in red, and IV estimates in blue. The instrument is the U.S. monetary policy shock from Bauer and Swanson (2023), interacted with country fixed effects, denoted $MShock_{n,t}$. Shocks are standardized to have a standard deviation of one. Standard errors are clustered at the country level, and 90% confidence intervals are reported. Panel (b) shows the first-stage F statistics across horizons, based on the robust procedure of Montiel Olea and Pflueger (2013). Panels (c) and (d) replicate the same analysis using the extended sample through 2023.

B. Phillips Multiplier with Nontradable Inflation



(a) Phillips Multiplier (2000-2019)

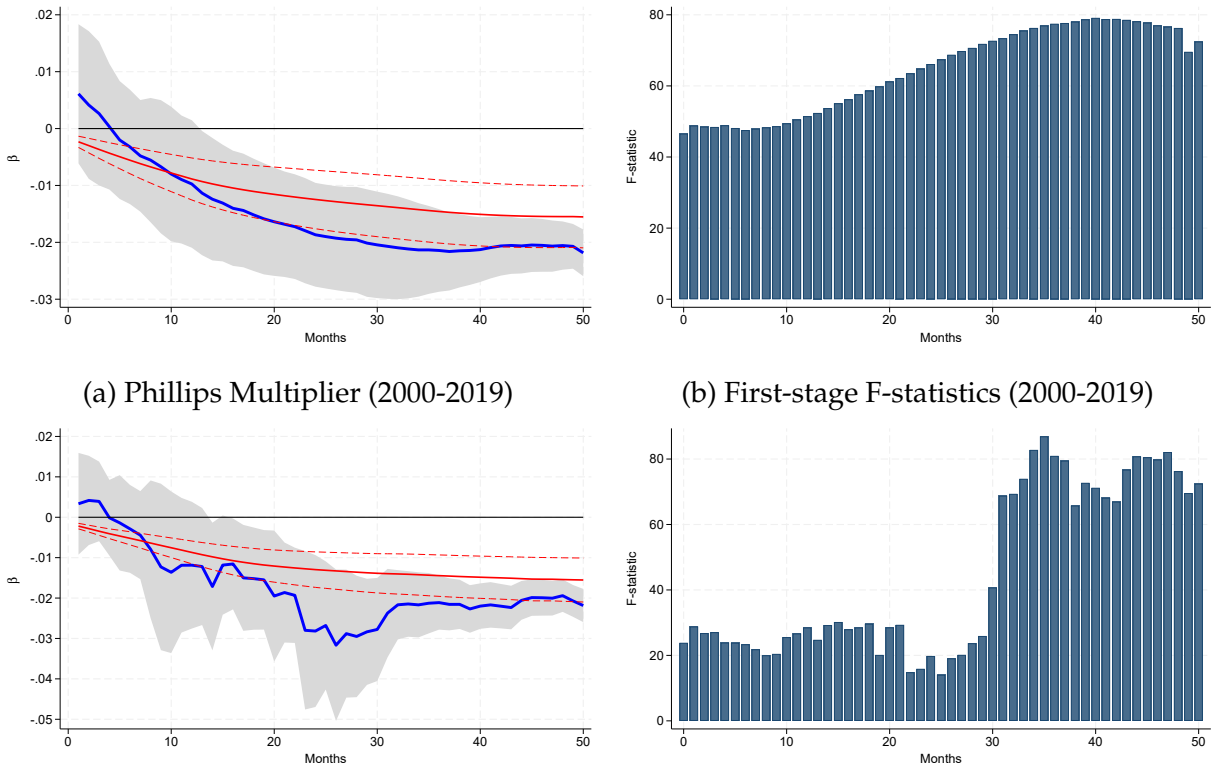
(b) First-stage F statistics (2000-2019)

(c) Phillips Multiplier (2000-2023)

(d) First-stage F statistics (2000-2023)

Figure A2: Phillips Multiplier for EA17 Countries with Nontradable (Core) Inflation

Notes: The figure presents results for the EA17 countries over two sample periods: 2000m1-2019m12 (Panels a and b) and 2000m1-2023m12 (Panels c and d). Panel (a) displays the Phillips multiplier estimates, β_h, from panel local projections (equation 2) over horizons $h = 1, \ldots, 50$ months. OLS estimates are shown in red, and IV estimates in blue. The instrument is the U.S. monetary policy shock from Bauer and Swanson (2023), interacted with country fixed effects, denoted $MShock_{n,t}$. Shocks are standardized to have a standard deviation of one. Standard errors are clustered at the country level, and 90% confidence intervals are reported. Panel (b) shows the first-stage F statistics across horizons, based on the robust procedure of Montiel Olea and Pflueger (2013). Panels (c) and (d) replicate the same analysis using the extended sample through 2023.



(a) Phillips Multiplier (2000-2019)

(b) First-stage F-statistics (2000-2019)

(c) Phillips Multiplier (2000-2023)

(d) First-stage F statistics (2000-2023)

Figure A3: Phillips Multiplier for EA12 Countries with Nontradable (Core) Inflation)

Notes: The figure presents results for the EA12 countries over two sample periods: 2000m1-2019m12 (Panels a and b) and 2000m1-2023m12 (Panels c and d). Panel (a) displays the Phillips multiplier estimates, β_h, from panel local projections (equation 2) over horizons $h = 1, \ldots, 50$ months. OLS estimates are shown in red, and IV estimates in blue. The instrument is the U.S. monetary policy shock from Bauer and Swanson (2023), interacted with country fixed effects, denoted $MShock_{n,t}$. Shocks are standardized to have a standard deviation of one. Standard errors are clustered at the country level, and 90% confidence intervals are reported. Panel (b) shows the first-stage F statistics across horizons, based on the robust procedure of Montiel Olea and Pflueger (2013). Panels (c) and (d) replicate the same analysis using the extended sample through 2023.

C. Forecast Robustness

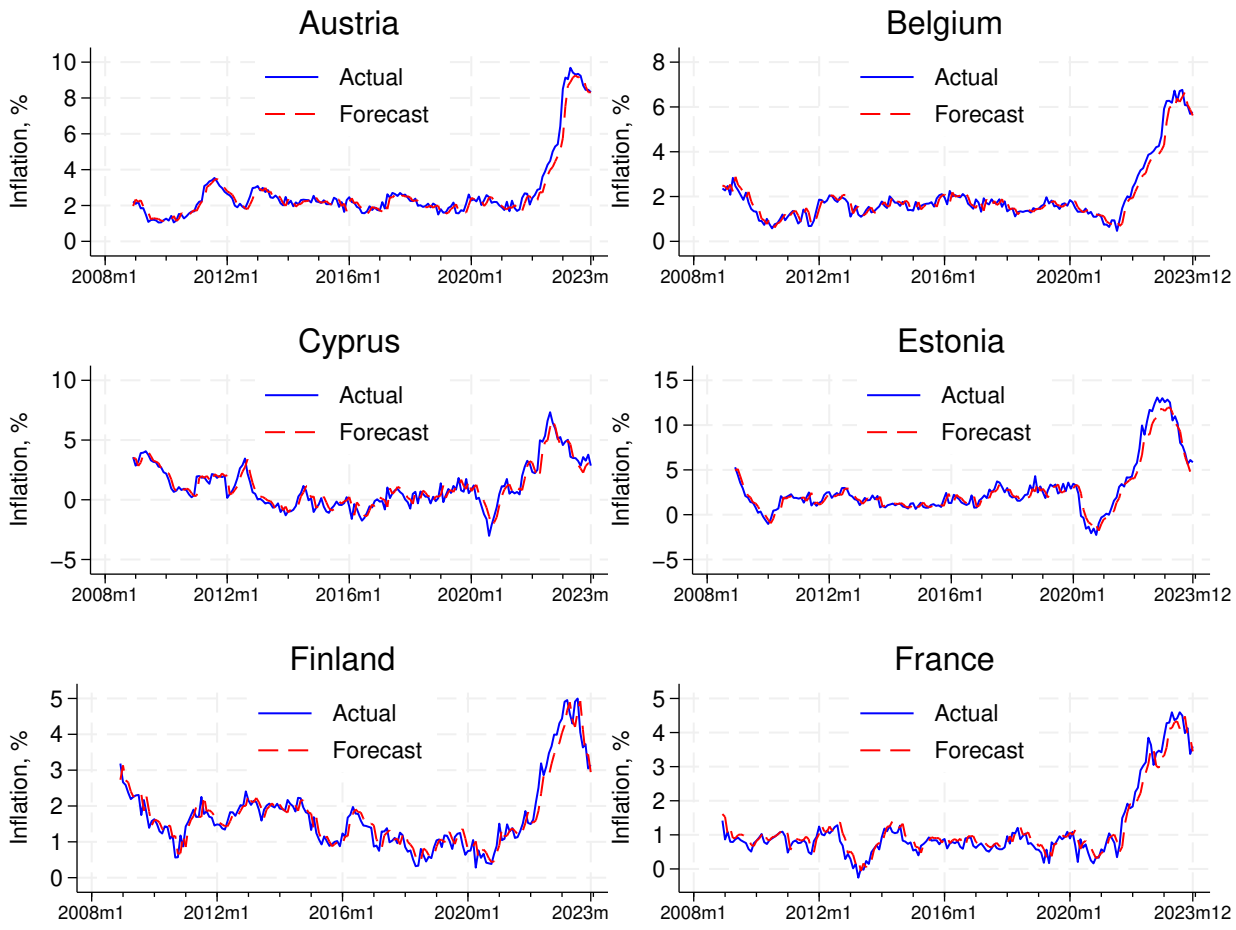


Figure A4: Panel Phillips curve forecasts vs. actual nontradables inflation 2008m:12-2023:m12

Notes: This figure plots nontradables inflation forecasts vs. actual inflation for 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

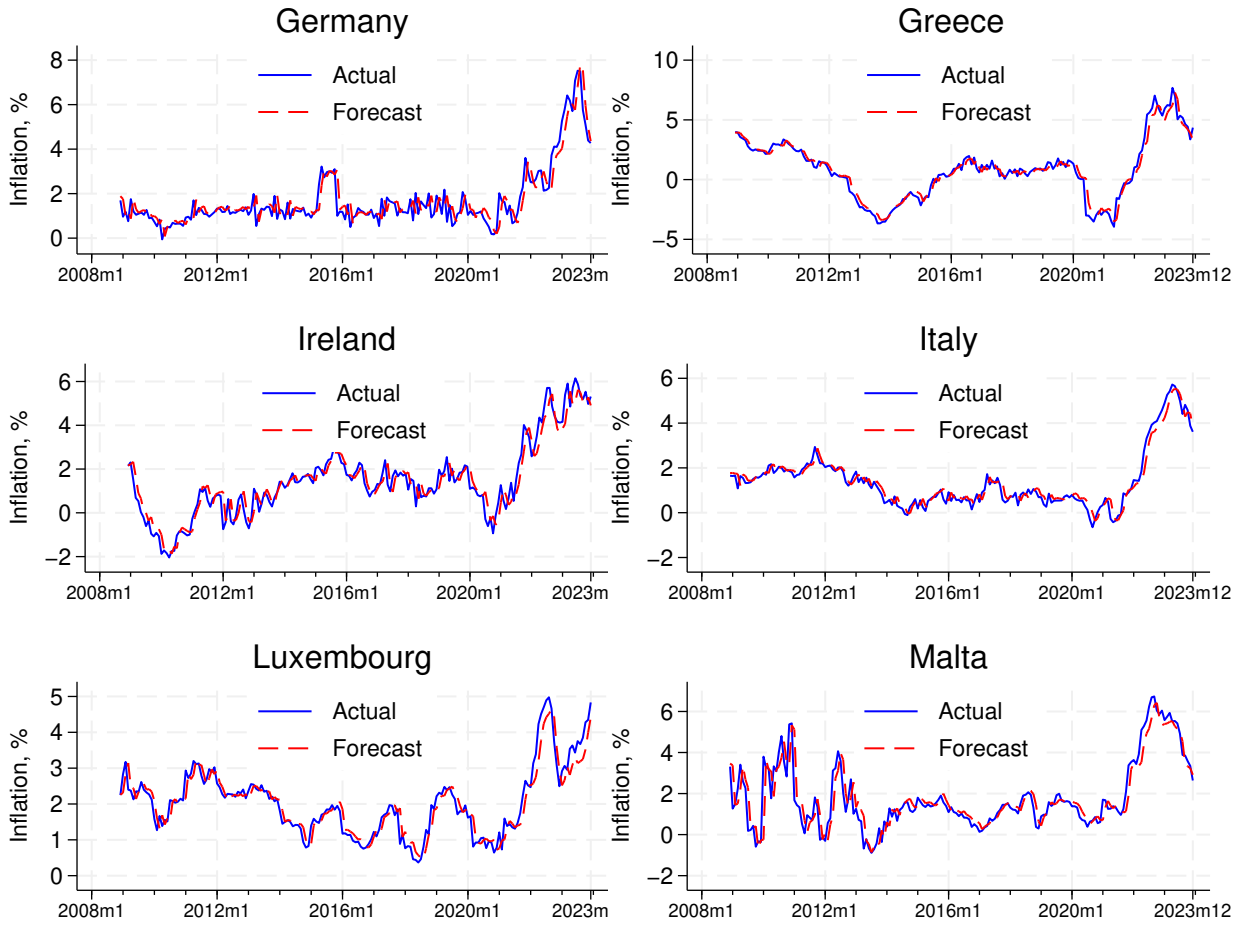


Figure A5: Panel Phillips curve forecasts vs. actual nontradables inflation 2008m:12-2023:m12

Notes: This figure plots nontradables inflation forecasts vs. actual inflation for 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

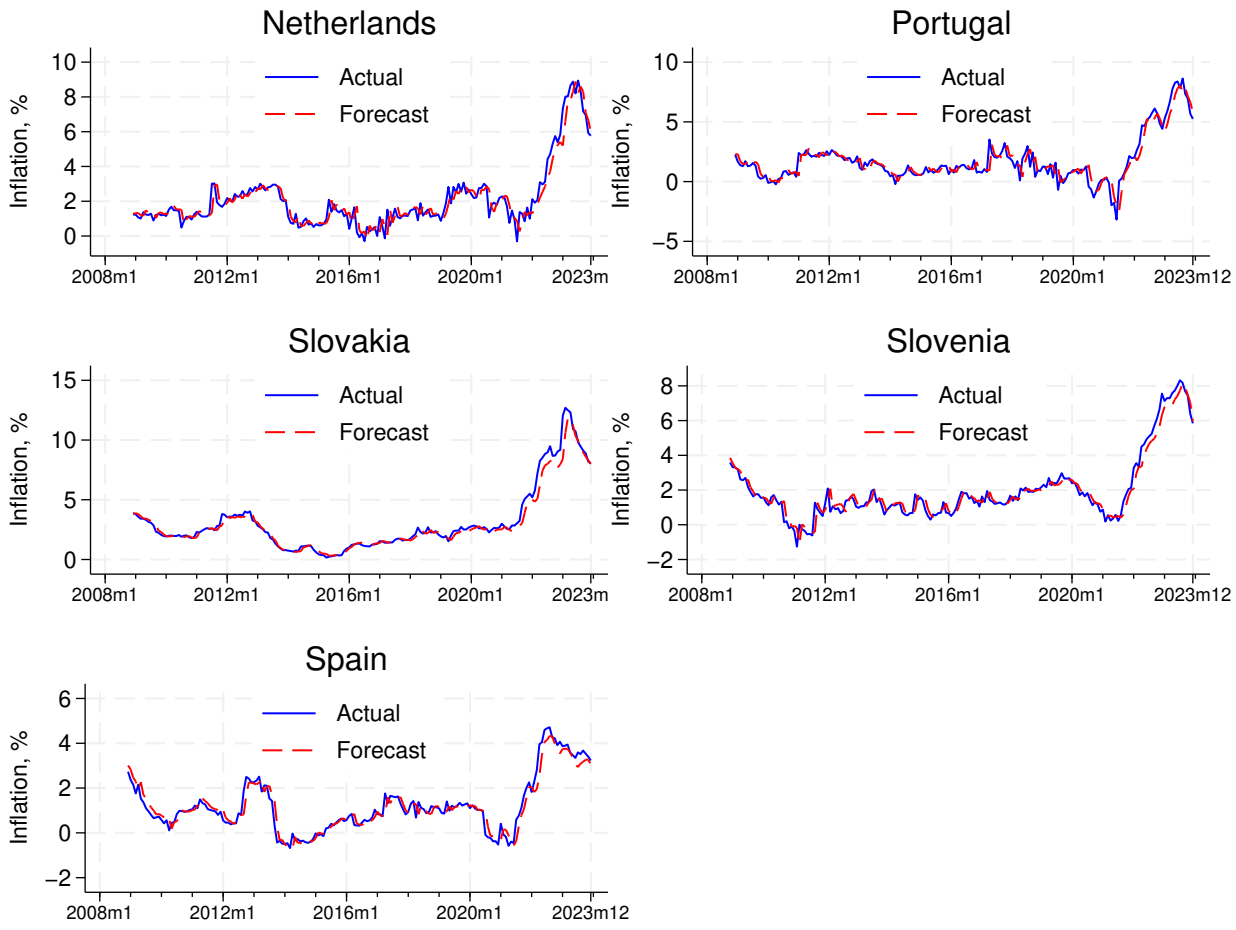


Figure A6: Panel Phillips curve forecasts vs. actual nontradables inflation 2008m:12-2023:m12

Notes: This figure plots nontradable inflation forecasts vs. actual inflation for 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

Table A1: Out-of-sample forecast performance of panel models with nontradable inflation for 2008m12:2019m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{17} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

		(a)			(b)			(c)		
Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value	Rel. RMSE Panel AR	DMW stat.	DMW p-value	Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.03	**0.71****	1.65	0.05	**0.73***	1.55	0.06	**0.73***	1.54	0.06
Belgium	0.02	1.03	-0.34	0.64	1.03	-0.39	0.65	1.04	-0.59	0.72
Cyprus	0.05	**0.92****	2.27	0.01	**0.92****	2.05	0.02	**0.93****	1.92	0.03
Estonia	0.05	**0.94*****	3.08	0.00	**0.94*****	3.33	0.00	**0.94*****	2.33	0.01
Finland	0.02	**0.97**	0.90	0.18	**0.96**	1.27	0.10	**0.96****	1.65	0.05
France	0.02	1.09	-1.63	0.95	1.08	-1.45	0.93	1.08	-1.39	0.92
Germany	0.04	1.07	-0.64	0.74	1.07	-0.64	0.74	1.07	-0.62	0.73
Greece	0.04	**0.76*****	3.83	0.00	**0.82*****	2.85	0.00	**0.82*****	2.81	0.00
Ireland	0.04	**0.94****	2.07	0.02	**0.94*****	2.56	0.01	**0.94****	2.24	0.01
Italy	0.02	**0.93****	2.22	0.01	**0.94****	1.91	0.03	**0.94****	1.86	0.03
Luxembourg	0.03	**0.87***	1.60	0.06	**0.87***	1.45	0.07	**0.88***	1.36	0.09
Malta	0.07	**0.96**	0.70	0.24	**0.96**	0.73	0.23	**0.97**	0.55	0.29
Netherlands	0.04	**0.98**	0.55	0.29	**0.98**	0.44	0.33	**0.98**	0.37	0.36
Portugal	0.05	1.02	-0.45	0.67	1.02	-0.53	0.70	1.03	-0.63	0.74
Slovakia	0.02	**0.75*****	2.83	0.00	**0.75*****	2.93	0.00	**0.78*****	2.57	0.01
Slovenia	0.04	**0.89*****	2.43	0.01	**0.90****	2.28	0.01	**0.89****	2.12	0.02
Spain	0.03	**0.83****	1.65	0.05	**0.83***	1.36	0.09	**0.84***	1.35	0.09

Notes: This table reports results for nontradable inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2019m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A2: Out-of-sample forecast performance of panel models with nontradable inflation for 2008m12:2023m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{17} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

Country	RMSE AR(12)	(a) Rel. RMSE Panel PC	DMW stat.	DMW p-value	(b) Rel. RMSE Panel AR	DMW stat.	DMW p-value	(c) Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.06	**0.54***	1.36	0.09	**0.53***	1.34	0.09	**0.52***	1.34	0.09
Belgium	0.03	**0.82**	1.16	0.12	**0.81**	1.13	0.13	**0.81**	1.11	0.13
Cyprus	0.07	**0.81***	1.52	0.06	**0.81***	1.50	0.07	**0.81***	1.54	0.06
Estonia	0.07	**0.91***	1.61	0.05	**0.88***	1.55	0.06	**0.88***	1.48	0.07
Finland	0.03	**0.94****	2.02	0.02	**0.93****	2.00	0.02	**0.93****	2.31	0.01
France	0.04	**0.53**	1.19	0.12	**0.52**	1.19	0.12	**0.53**	1.18	0.12
Germany	0.06	**0.77**	1.19	0.12	**0.77**	1.18	0.12	**0.77**	1.18	0.12
Greece	0.05	**0.90*****	2.34	0.01	**0.92****	1.85	0.03	**0.93****	1.62	0.05
Ireland	0.05	**0.81***	1.43	0.08	**0.80***	1.41	0.08	**0.81***	1.39	0.08
Italy	0.04	**0.73***	1.44	0.08	**0.73***	1.40	0.08	**0.72***	1.40	0.08
Luxembourg	0.03	**0.87****	1.93	0.03	**0.87****	1.80	0.04	**0.86****	1.82	0.03
Malta	0.07	**0.91***	1.50	0.07	**0.90***	1.51	0.07	**0.91***	1.36	0.09
Netherlands	0.05	**0.88***	1.36	0.09	**0.88**	1.26	0.10	**0.89**	1.12	0.13
Portugal	0.07	**0.85***	1.55	0.06	**0.85***	1.49	0.07	**0.85***	1.44	0.08
Slovakia	0.04	1.04	-0.21	0.58	0.94	**0.36**	0.36	0.95	**0.30**	0.38
Slovenia	0.04	**0.90*****	2.47	0.01	**0.89*****	2.36	0.01	**0.89****	2.30	0.01
Spain	0.03	**0.86****	1.87	0.03	**0.84****	1.80	0.04	**0.84****	1.78	0.04

Notes: This table reports results for nontradable inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2023m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A3: Out-of-sample forecast performance of panel models with headline inflation for 2009m6:2023m12 (h=18)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{17} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

		(a)			(b)			(c)		
Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value	Rel. RMSE Panel AR	DMW stat.	DMW p-value	Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.04	**0.89****	1.74	0.04	**0.85****	1.74	0.04	**0.85****	1.67	0.05
Belgium	0.08	**0.76****	1.96	0.02	**0.75****	1.96	0.02	**0.75****	1.96	0.02
Cyprus	0.07	**0.81*****	2.67	0.00	**0.81*****	2.67	0.00	**0.81*****	2.64	0.00
Estonia	0.08	1.00	-0.04	0.52	**0.96**	0.36	0.36	**0.96**	0.37	0.36
Finland	0.04	**0.89*****	2.49	0.01	**0.87*****	2.40	0.01	**0.87****	2.32	0.01
France	0.04	**0.83***	1.44	0.08	**0.80****	1.66	0.05	**0.80***	1.64	0.05
Germany	0.06	**0.84***	1.40	0.08	**0.83***	1.42	0.08	**0.83***	1.39	0.08
Greece	0.07	**0.80****	2.05	0.02	**0.79****	2.17	0.02	**0.80****	2.02	0.02
Ireland	0.06	**0.64***	1.47	0.07	**0.62***	1.49	0.07	**0.62***	1.47	0.07
Italy	0.05	**0.87****	1.73	0.04	**0.85***	1.58	0.06	**0.86***	1.56	0.06
Luxembourg	0.07	**0.80*****	2.46	0.01	**0.80*****	2.48	0.01	**0.79*****	2.41	0.01
Malta	0.05	**0.79****	2.06	0.02	**0.79****	2.07	0.02	**0.79****	2.02	0.02
Netherlands	0.08	**0.93***	1.46	0.07	**0.92***	1.50	0.07	**0.93***	1.50	0.07
Portugal	0.05	**0.82***	1.42	0.08	**0.81***	1.41	0.08	**0.82***	1.39	0.08
Slovakia	0.04	1.26	-1.08	0.86	1.17	-0.92	0.82	1.16	-0.92	0.82
Slovenia	0.05	**0.98**	0.56	0.29	**0.96***	1.29	0.10	**0.96**	0.99	0.16
Spain	0.07	**0.77****	1.83	0.03	**0.76****	1.94	0.03	**0.75****	1.97	0.02

Notes: This table reports results for headline inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=18 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2023m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A4: Out-of-sample forecast performance of panel models with headline inflation for 2009m12:2023m12 (h=24)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{17} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

Country	(a) RMSE AR(12)	(a) Rel. RMSE Panel PC	(a) DMW stat.	(a) DMW p-value	(b) Rel. RMSE Panel AR	(b) DMW stat.	(b) DMW p-value	(c) Rel. RMSE Panel IRF	(c) DMW stat.	(c) DMW p-value
Austria	0.04	**0.92****	1.87	0.03	**0.89****	1.99	0.02	**0.88****	1.91	0.03
Belgium	0.08	**0.79****	2.09	0.02	**0.79****	2.08	0.02	**0.79****	2.07	0.02
Cyprus	0.07	**0.80*****	3.28	0.00	**0.80*****	3.29	0.00	**0.80*****	3.25	0.00
Estonia	0.08	**0.98**	0.20	0.42	**0.95**	0.46	0.32	**0.96**	0.47	0.32
Finland	0.04	**0.94*****	2.40	0.01	**0.93*****	3.05	0.00	**0.93*****	3.01	0.00
France	0.03	**0.89****	1.97	0.03	**0.87****	2.15	0.02	**0.87****	2.08	0.02
Germany	0.06	**0.83***	1.62	0.05	**0.82***	1.64	0.05	**0.82***	1.63	0.05
Greece	0.08	**0.73****	2.16	0.02	**0.73****	2.21	0.01	**0.74****	2.13	0.02
Ireland	0.06	**0.72***	1.49	0.07	**0.70***	1.51	0.07	**0.71***	1.47	0.07
Italy	0.05	**0.91****	2.01	0.02	**0.90***	1.82	0.04	**0.90***	1.77	0.04
Luxembourg	0.07	**0.80****	2.92	0.00	**0.79****	2.94	0.00	**0.80****	2.74	0.00
Malta	0.05	**0.76****	2.38	0.01	**0.75****	2.39	0.01	**0.76****	2.33	0.01
Netherlands	0.07	**0.97**	1.21	0.11	**0.98**	1.11	0.13	**0.98**	1.08	0.14
Portugal	0.05	**0.82**	1.21	0.11	**0.82**	1.21	0.11	**0.82**	1.17	0.12
Slovakia	0.04	1.24	-1.19	0.88	1.16	-0.99	0.84	1.16	-0.99	0.84
Slovenia	0.05	**0.95**	1.17	0.12	**0.94****	1.77	0.04	**0.94***	1.56	0.06
Spain	0.07	**0.81****	1.68	0.05	**0.80****	1.76	0.04	**0.80****	1.76	0.04

Notes: This table reports results for headline inflation forecasts from 17 Euro Area member countries. The forecast horizon is h=24 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2023m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A5: Out-of-sample forecast performance of panel models for EA12 with headline inflation for 2008m12:2019m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{12} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

		(a)			(b)			(c)		
Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value	Rel. RMSE Panel AR	DMW stat.	DMW p-value	Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.03	**0.92**	1.26	0.10	**0.92***	1.32	0.09	**0.93**	1.13	0.13
Belgium	0.04	**0.95**	0.71	0.24	**0.95**	0.72	0.24	**0.95**	0.73	0.23
Finland	0.03	**0.90*****	3.02	0.00	**0.89*****	3.04	0.00	**0.89*****	3.07	0.00
France	0.03	**0.71**	1.23	0.11	**0.71**	1.23	0.11	**0.71**	1.23	0.11
Germany	0.03	1.07	-1.74	0.96	1.06	-1.48	0.93	1.06	-1.50	0.93
Greece	0.06	**0.78****	1.75	0.04	**0.80****	1.67	0.05	**0.80***	1.63	0.05
Ireland	0.03	**0.94**	1.06	0.15	**0.95**	0.94	0.18	**0.96**	0.80	0.21
Italy	0.03	**0.82****	2.08	0.02	**0.82****	2.06	0.02	**0.82****	1.99	0.02
Luxembourg	0.05	**0.88****	1.83	0.03	**0.88****	1.81	0.03	**0.88****	1.86	0.03
Netherlands	0.03	**0.96**	0.98	0.16	**0.96**	1.04	0.15	**0.96**	0.93	0.18
Portugal	0.04	1.00	-0.11	0.54	1.01	-0.24	0.59	1.01	-0.37	0.64
Spain	0.06	**0.65***	1.41	0.08	**0.66***	1.39	0.08	**0.66***	1.39	0.08

Notes: This table reports results for headline inflation forecasts from 12 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2019m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A6: Out-of-sample forecast performance of panel models with headline inflation for 2008m12:2023m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(b) Panel AR(12): $\pi_{i,t+h} = \alpha_0 + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
(c) Panel IRF: $\pi_{i,t+h} = \alpha_0 + \sum_{n=1}^{12} \beta_i^{\pi} MShock_{i,t}^{n} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
relative to:
(d) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

		(a)			(b)			(c)		
Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value	Rel. RMSE Panel AR	DMW stat.	DMW p-value	Rel. RMSE Panel IRF	DMW stat.	DMW p-value
Austria	0.05	**0.93***	1.31	0.10	**0.89***	1.65	0.05	**0.89***	1.61	0.05
Belgium	0.07	**0.85***	1.52	0.06	**0.84****	1.59	0.06	**0.83***	1.59	0.06
Finland	0.04	**0.92****	1.92	0.03	**0.89*****	2.57	0.00	**0.89*****	2.51	0.01
France	0.04	**0.87**	0.99	0.16	**0.84**	1.22	0.11	**0.85**	1.17	0.12
Germany	0.05	**0.90**	1.14	0.13	**0.88**	1.18	0.12	**0.88**	1.14	0.13
Greece	0.07	**0.88****	1.76	0.04	**0.87****	2.04	0.02	**0.87****	1.92	0.03
Ireland	0.06	**0.76***	1.64	0.05	**0.74***	1.62	0.05	**0.75***	1.58	0.06
Italy	0.06	**0.87*****	2.47	0.01	**0.85****	2.25	0.01	**0.85****	2.23	0.01
Luxembourg	0.07	**0.89****	1.74	0.04	**0.87****	1.97	0.02	**0.87****	1.96	0.03
Netherlands	0.08	**0.98**	0.32	0.37	**0.97**	0.46	0.33	**0.97**	0.42	0.34
Portugal	0.06	**0.83***	1.62	0.05	**0.82***	1.58	0.06	**0.82***	1.55	0.06
Spain	0.07	**0.82***	1.38	0.08	**0.78***	1.62	0.05	**0.78****	1.65	0.05

Notes: This table reports results for headline inflation forecasts from 12 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2023m12. We report: (a) the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark; (b) the relative RMSE of the Panel AR(12) model versus the AR(12) benchmark; and (c) the relative RMSE of the Panel IRF model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the respective panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the given panel model outperforms the AR(12) benchmark. Standard errors are adjusted for cross-sectional dependence following Driscoll and Kraay (1998). ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A7: Out-of-sample forecast performance of individual country Phillips curves (h=12)
(a) Phillips curve: $\pi_{t+h} = \alpha_0 + \beta u_t + \sum_{j=1}^{12} \beta_j \pi_{t-j} + \varepsilon_{t+h}$
(b) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

	(a) Headline 2019			**(b) Headline 2023**			**(c) Nontradable 2019**			**(d) Nontradable 2023**		
Country	Rel. RMSE PC/AR(12)	CW stat.	CW p-value	Rel. RMSE PC/AR(12)	CW stat.	CW p-value	Rel. RMSE PC/AR(12)	CW stat.	CW p-value	Rel. RMSE PC/AR(12)	CW stat.	CW p-value
Austria	1.00***	3.27	0.00	1.03***	3.05	0.00	1.01*	1.51	0.07	**0.97***	1.45	0.07
Belgium	1.04*	1.61	0.05	1.07	1.25	0.11	1.01***	4.06	0.00	1.11**	1.87	0.03
Cyprus	1.12	0.24	0.40	1.10	0.55	0.29	1.06	0.98	0.16	1.20	-0.71	0.76
Estonia	1.14	-0.24	0.59	1.07	0.49	0.31	1.11	-0.59	0.72	1.34	-1.18	0.88
Finland	1.09	0.61	0.27	1.05	1.28	0.10	1.02**	1.83	0.03	1.05	1.12	0.13
France	1.11	0.79	0.21	1.09	0.71	0.24	0.99***	4.01	0.00	1.01*	1.32	0.09
Germany	1.00***	5.29	0.00	1.07*	1.36	0.09	1.01**	2.96	0.00	1.01**	1.99	0.02
Greece	1.08**	1.84	0.03	1.05**	1.83	0.03	1.07*	1.30	0.10	1.01**	1.99	0.02
Ireland	1.39	-1.44	0.93	1.13	0.35	0.36	1.15	-0.97	0.83	1.07	0.92	0.18
Italy	1.11**	2.08	0.02	1.03*	1.64	0.05	1.29	-1.19	0.88	1.10	0.53	0.30
Luxembourg	**0.96****	2.40	0.01	1.03*	1.48	0.07	1.07*	1.38	0.08	1.01**	2.05	0.02
Malta	1.03**	2.39	0.01	1.05***	2.68	0.00	1.01**	1.87	0.03	1.01***	2.49	0.01
Netherlands	1.03*	1.89	0.03	1.14	-0.84	0.80	1.02**	2.12	0.02	1.06	1.27	0.10
Portugal	1.30	-1.41	0.92	1.13	-0.59	0.72	1.07	1.09	0.14	1.03**	2.29	0.01
Slovakia	1.12	-0.64	0.74	1.19	-0.87	0.81	1.28	-0.67	0.75	1.23	0.28	0.39
Slovenia	1.16	-1.45	0.93	1.11	0.04	0.48	1.01*	1.48	0.07	1.23	-0.87	0.81
Spain	**0.93***	1.13	0.13	**0.98**	1.27	0.10	1.34	-0.77	0.78	1.40	-1.40	0.92

Notes: This table reports results for headline and nontradable inflation forecasts from each of 17 Euro Area member countries. The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12. The forecast period ends in 2019m2 or 2023m12 as indicated in sections (a)-(d). We report the relative RMSE of the individual-country Phillips Curve (PC) model versus the AR(12) benchmark. A relative RMSE less than one indicates that the PC model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Clark and West (2007) (CW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy between the models; the alternative hypothesis implies that the individual-country PC model outperforms the AR(12) benchmark. Newey-West standard errors are used to adjust for serial correlation (12 lags) and heteroskedasticity. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A8: Out-of-sample forecast performance of panel models with headline inflation for 2008m12:2023m12 (h=12)
(a) Phillips curve (PC): $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$
Relative to:
(b) AR(12): $\pi_{t+h} = \gamma_0 + \sum_{j=1}^{12} \gamma_j \pi_{t-j} + \varepsilon_{t+h}$ (benchmark)

Country	RMSE AR(12)	Rel. RMSE Panel PC	DMW stat.	DMW p-value
Austria	0.05	**0.88***	1.62	0.05
Belgium	0.07	**0.83****	1.70	0.05
Cyprus	0.07	**0.85****	1.97	0.03
Estonia	0.08	1.01	–0.10	0.54
Finland	0.04	**0.90****	2.23	0.01
France	0.03	**0.89**	1.08	0.14
Germany	0.05	**0.90**	1.14	0.13
Greece	0.07	**0.86****	2.34	0.01
Ireland	0.06	**0.73***	1.62	0.05
Italy	0.06	**0.84****	2.32	0.01
Luxembourg	0.07	**0.86****	2.19	0.01
Malta	0.05	**0.83****	1.76	0.04
Netherlands	0.08	**0.96**	0.51	0.31
Portugal	0.06	**0.81***	1.52	0.06
Slovakia	0.04	1.11	–0.80	0.79
Slovenia	0.05	**0.97**	0.59	0.28
Spain	0.07	**0.78***	1.65	0.05

Notes: This table reports results for headline inflation forecasts from 16 Euro Area member countries (EA17 excluding Germany). The forecast horizon is $h = 12$ months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12 for all countries with the following exceptions: Cyprus starts in 2000:m12, France in 2003:m12, Estonia in 2001:m1. The forecast period ends in 2023m12. We report the relative RMSE of the Panel Phillips Curve (PC) model versus the AR(12) benchmark. The first column presents the absolute RMSE of the AR(12) benchmark. A relative RMSE less than one indicates that the panel model provides more accurate forecasts than the AR(12) benchmark. To assess statistical significance, we report the Diebold and Mariano (1995) and West (1996) (DMW) test statistic and its corresponding p-value. Inference is based on standard normal critical values for a one-sided test. The null hypothesis is equal forecast accuracy; the alternative implies that the panel model outperforms the benchmark. Standard errors are adjusted for cross-sectional dependence (Driscoll and Kraay, 1998). ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

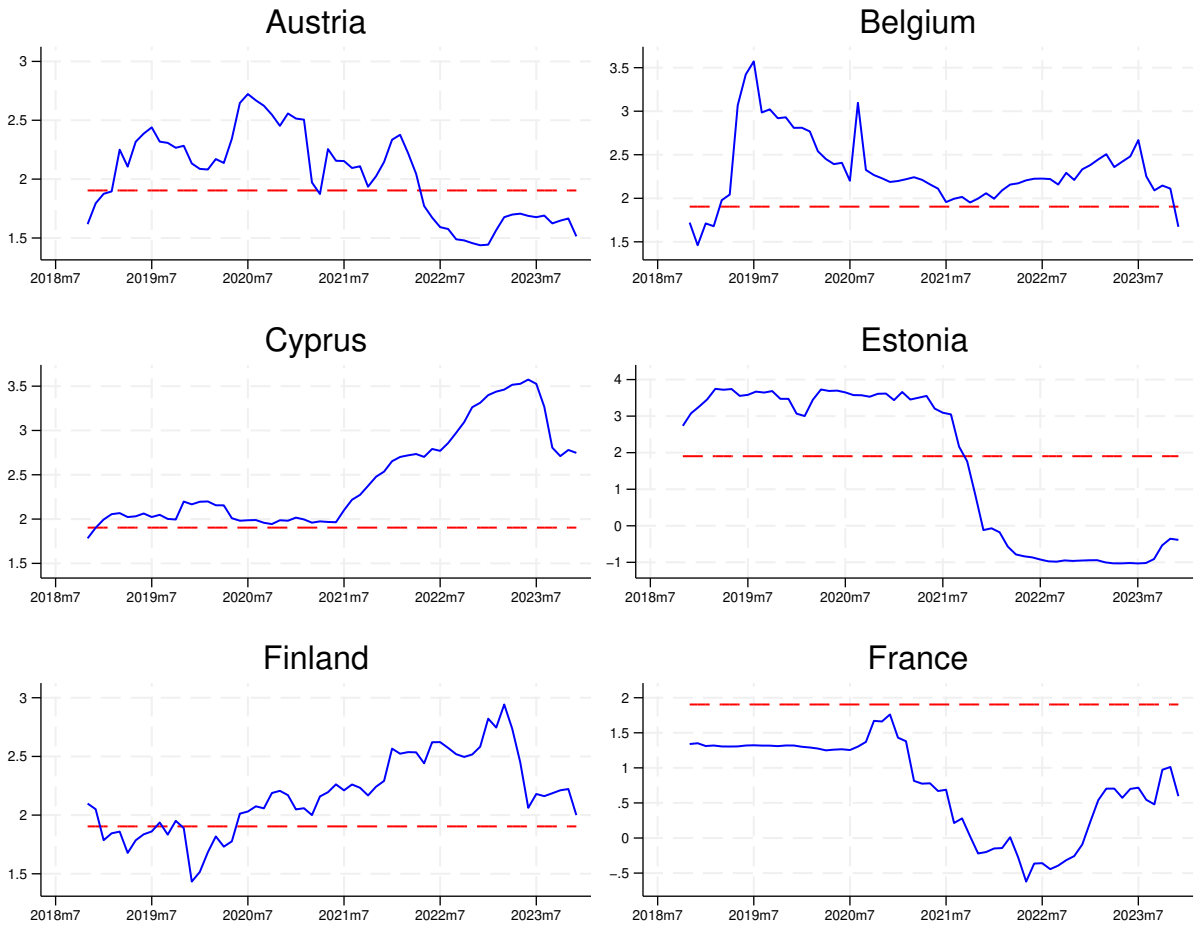


Figure A7: Giacomini-Rossi (2010) Fluctuation Test

Notes: This figure plots the results from Giacomini and Rossi (2010) fluctuation test, evaluating the performance of headline inflation forecasts from the benchmark AR(12) model and the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$. The solid blue line plots the Giacomini-Rossi test statistic. The dashed red line plots the critical value of a one-sided test at 10% significance level. A test statistic greater than zero suggests the panel PC model outperforms the AR(12) benchmark for a given country. A test statistic above the critical value suggests the panel PC is significantly better. The dates on the horizontal axis refer to the end point of each fluctuation test window, where the two forecast models are compared. Each test window covers a 10-year period of monthly forecasts; for example, the test statistic at 2019m7 is based on forecasts from the preceding 10 years, spanning 2009m8 to 2019m7.

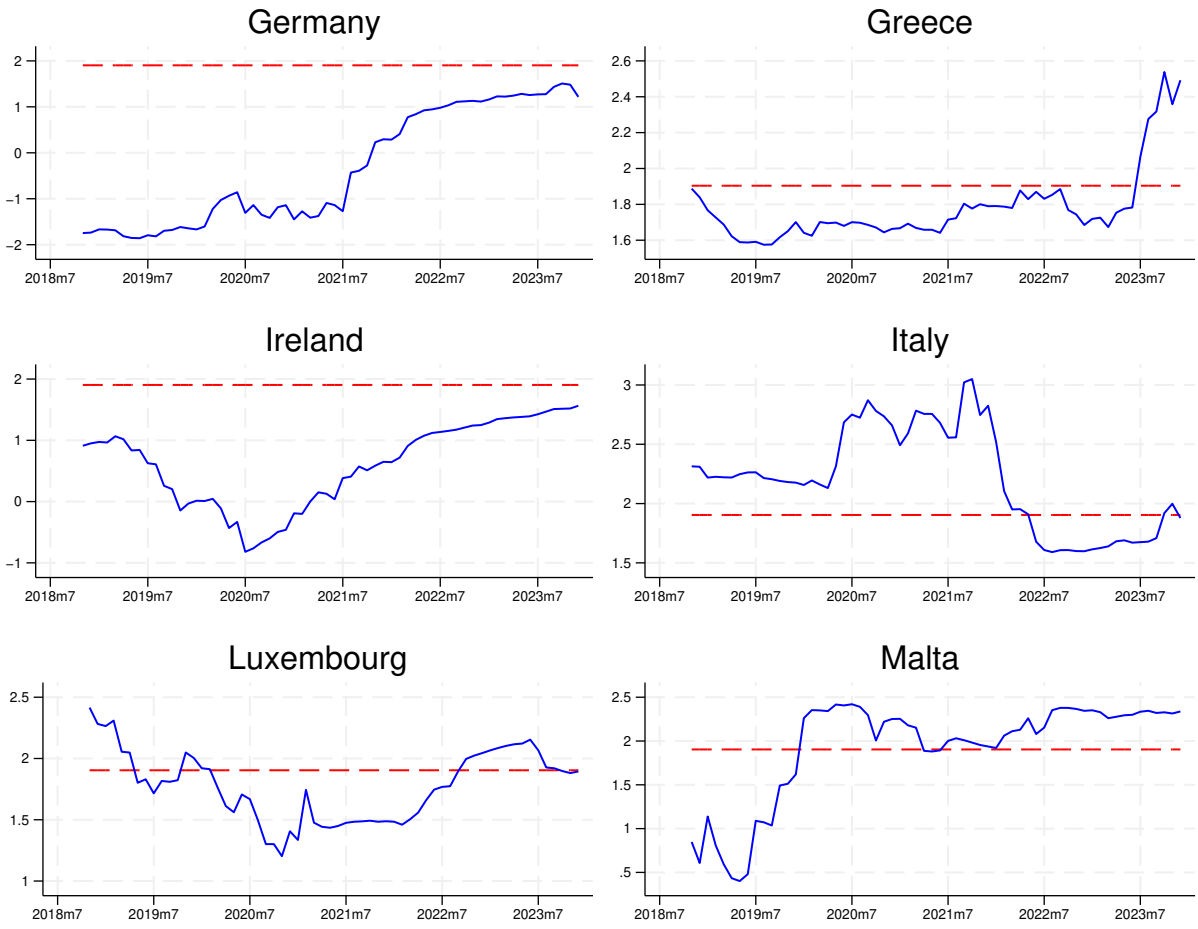


Figure A8: Giacomini-Rossi (2010) Fluctuation Test

Notes: This figure plots the results from Giacomini and Rossi (2010) fluctuation test, evaluating the performance of headline inflation forecasts from the benchmark AR(12) model and the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$. The solid blue line plots the Giacomini-Rossi test statistic. The dashed red line plots the critical value of a one-sided test at 10% significance level. A test statistic greater than zero suggests the panel PC model outperforms the AR(12) benchmark for a given country. The dates on the horizontal axis refer to the end point of each fluctuation test window, where the two forecast models are compared. Each test window covers a 10-year period of monthly forecasts; for example, the test statistic at 2019m7 is based on forecasts from the preceding 10 years, spanning 2009m8 to 2019m7.

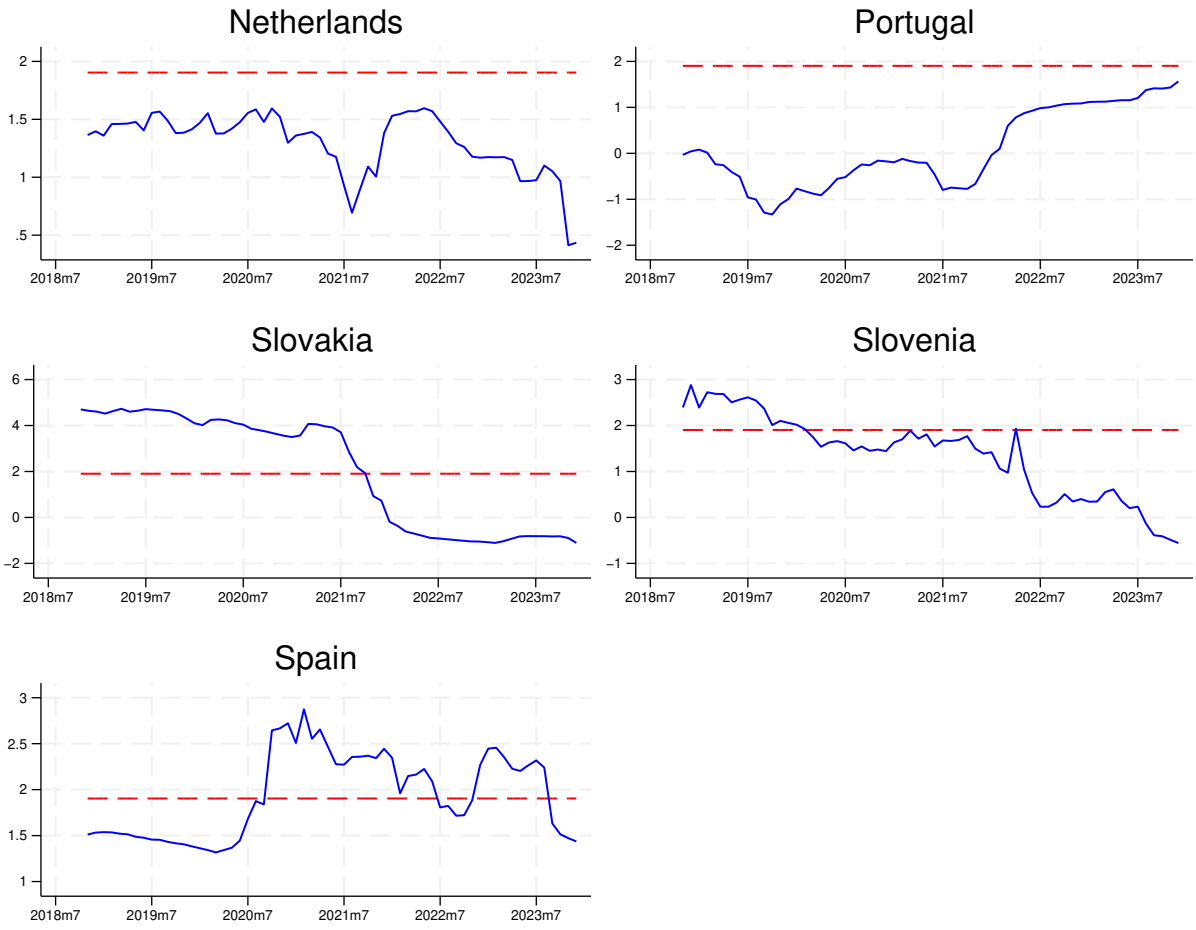


Figure A9: Giacomini-Rossi (2010) Fluctuation Test

Notes: This figure plots the results from Giacomini and Rossi (2010) fluctuation test, evaluating the performance of headline inflation forecasts from the benchmark AR(12) model and the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$. The solid blue line plots the Giacomini-Rossi test statistic. The dashed red line plots the critical value of a one-sided test at 10% significance level. A test statistic greater than zero suggests the panel PC model outperforms the AR(12) benchmark for a given country. A test statistic above the critical value suggests the panel PC is significantly better. The dates on the horizontal axis refer to the end point of each fluctuation test window, where the two forecast models are compared. Each test window covers a 10-year period of monthly forecasts; for example, the test statistic at 2019m7 is based on forecasts from the preceding 10 years, spanning 2009m8 to 2019m7.

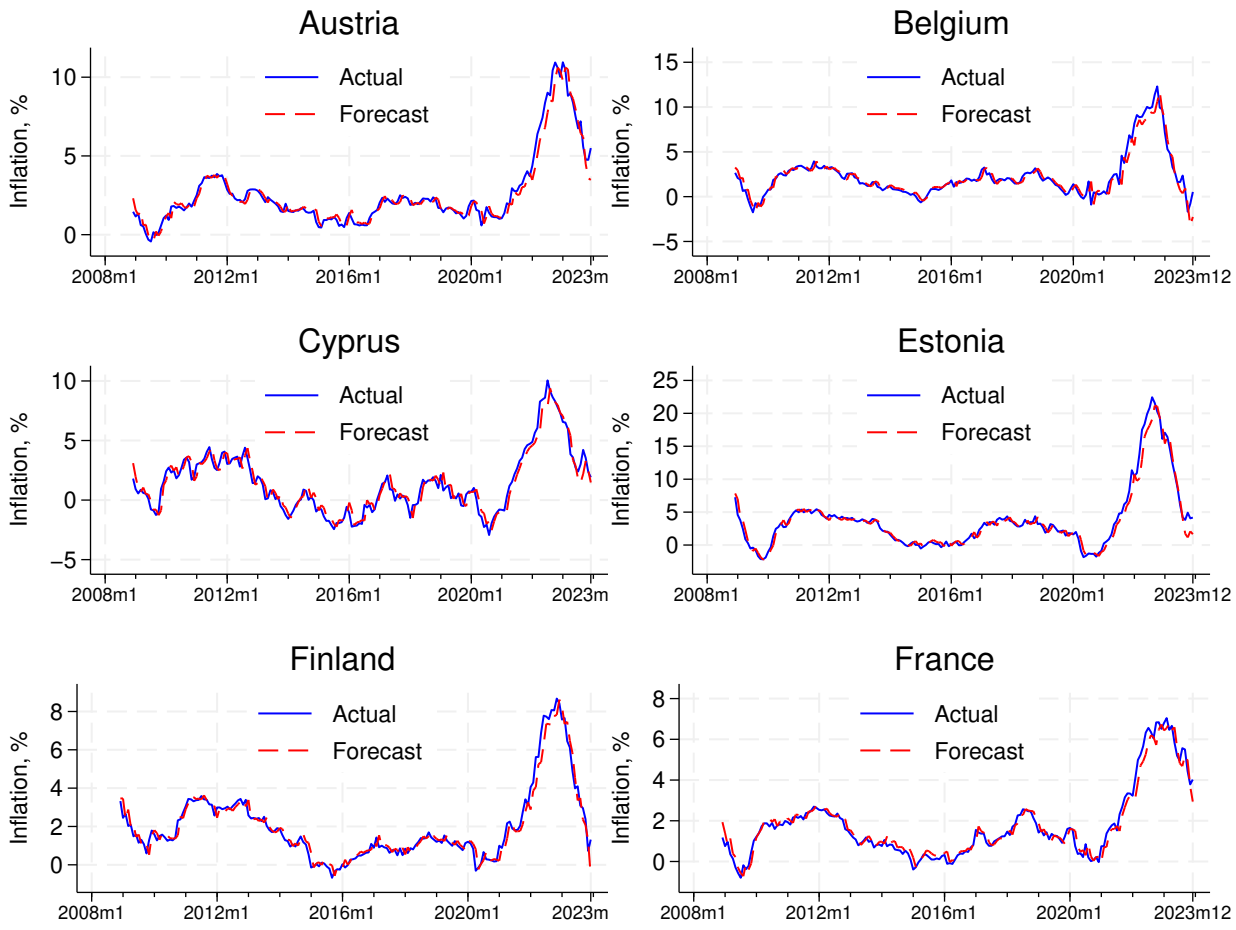


Figure A10: Panel Phillips curve forecasts vs. actual headline inflation 2008m:12-2023:m12

Notes: This figure plots headline inflation forecasts vs. actual inflation for 16 Euro Area member countries (EA17 excluding Germany). The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12 for all countries with the following exceptions: Cyprus starts in 2000:m12, France in 2003:m12, Estonia in 2001:m1. The forecast period ends in 2023m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

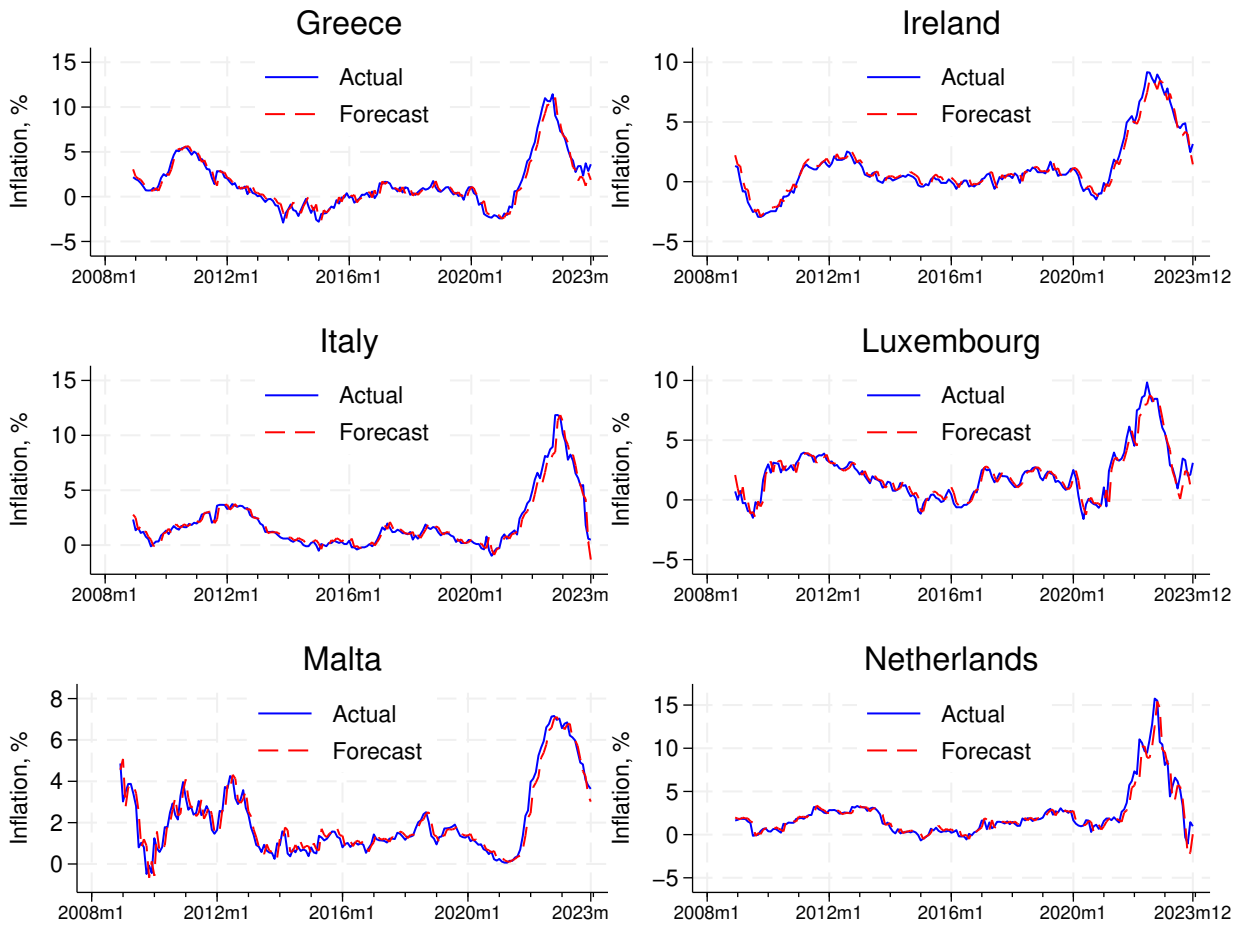


Figure A11: Panel Phillips curve forecasts vs. actual headline inflation 2008m:12-2023:m12

Notes: This figure plots headline inflation forecasts vs. actual inflation for 16 Euro Area member countries (EA17 excluding Germany). The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12 for all countries with the following exceptions: Cyprus starts in 2000:m12, France in 2003:m12, Estonia in 2001:m1. The forecast period ends in 2023m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

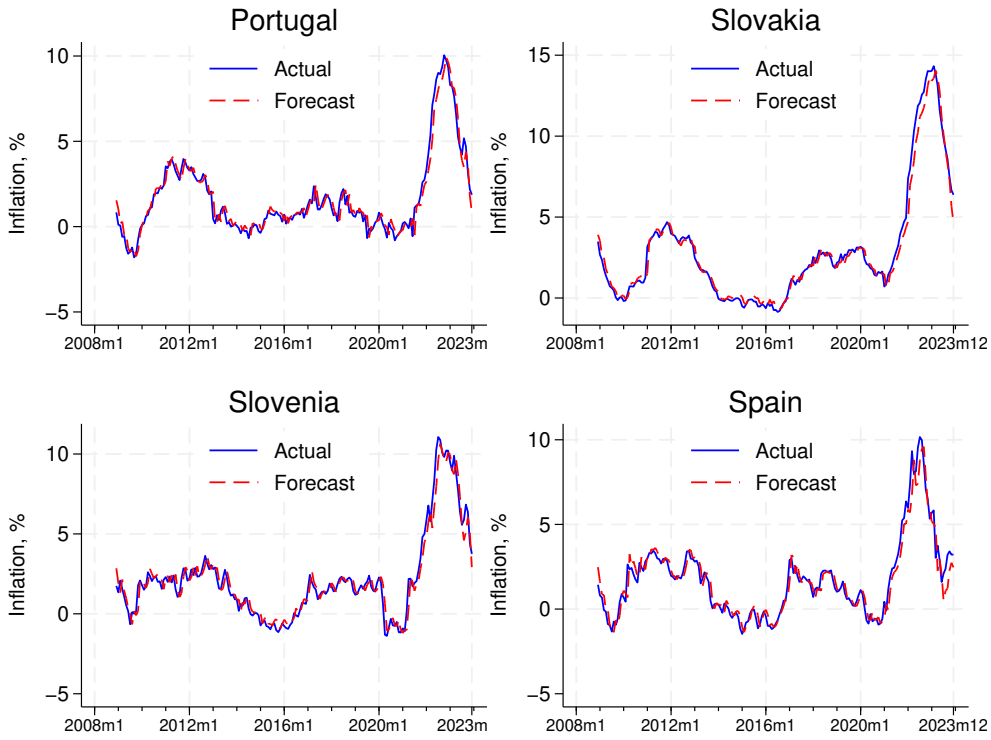


Figure A12: Panel Phillips curve forecasts vs. actual headline inflation 2008m:12-2023:m12

Notes: This figure plots headline inflation forecasts vs. actual inflation for 16 Euro Area member countries (EA17 excluding Germany). The forecast horizon is h=12 months. Rolling window forecasts are generated using a window length of 96 months, with the initial estimation window covering 2000m1-2007m12 for all countries with the following exceptions: Cyprus starts in 2000:m12, France in 2003:m12, Estonia in 2001:m1. The forecast period ends in 2023m12. The forecast period starts in 2008m12 and ends in 2023m12. The forecasts are based on the Panel Phillips Curve (PC) model: $\pi_{i,t+h} = \alpha_0 + \beta u_{it} + \sum_{j=1}^{12} \beta_j \pi_{i,t-j} + \varepsilon_{i,t+h}$.

D. Data Details

Tradable and Non-tradable Price Data Construction

The Harmonized Index of Consumer Prices (HICP) is available from Eurostat at the 4-digit COICOP level with monthly frequency. We construct tradable and non-tradable price indices based on 12 COICOP categories. The weights for each category are provided for each EU country and updated annually. Thus, tradable and non-tradable price indices are constructed as weighted averages of the relevant categories for each country over time. These series are used throughout the paper.

COICOP Code	Category	Tradability	Weights
CP01	Food and Non-alcoholic Beverages	Tradable	16.8
CP02	Alcoholic Beverages and Tobacco	Tradable	4.2
CP03	Clothing and Footwear	Tradable	5.0
CP04	Housing, Water, Electricity, Gas and Other Fuels		
CP041	Actual Rentals for Housing	Nontradable	5.1
CP043	Maintenance and Repair of the Dwelling	Nontradable	1.4
CP044	Water Supply and Miscellaneous Services	Nontradable	2.3
CP045	Electricity, Gas and Other Fuels	Tradable	6.0
CP05	Furnishings, Household Equipment and Maintenance	Tradable	6.3
CP06	Health	Nontradable	5.3
CP07	Transportation		
CP071	Purchases of Vehicles	Tradable	3.7
CP072	Operation of Personal Transport Equipment	Tradable	8.6
CP073	Transport Services	Nontradable	2.5
CP08	Communications	Nontradable	2.7
CP09	Recreation and Culture	Nontradable	8.9
CP10	Education	Nontradable	1.0
CP11	Restaurants and Hotels	Nontradable	10.3
CP12	Miscellaneous Goods and Services	Nontradable	9.7

Table A9: COICOP Categories and Tradability Classification

Notes: Weights are percentages reflecting the expenditure shares of each category in the Euro Area, in 2024.

Hazell et al. (2022) categorize entry-level items (ELIs) into tradable and non-tradable, where ELIs related to housing services, health, recreational activities, education, and restaurants are considered non-tradable. Although there is no one-to-one mapping between ELIs and COICOP categories, we adopt similar logic by categorizing food (CP01), alcoholic beverages (CP02), clothing (CP03), only the electricity and gas component of housing costs (CP045), furnishings (CP05), and transportation (CP071 and CP072) as tradable. The remaining categories are considered non-tradable.

Aggregate Inflation and Unemployment Data

Data Source	Variable Code	Definition
Eurostat	PRC_HICP_MIDX (CP00)	All Items - HICP
Eurostat	PRC_HICP_MIDX (TOT_X_NRG_FOOD)	Overall index excluding energy, food, alcohol and tobacco
Eurostat	UNE_RT_M (NSA)	Unemployment Rate (not seasonally adjusted)
Eurostat	UNE_RT_M (SA)	Unemployment Rate (seasonally adjusted)
Eurostat	InflationExpectation	HICP Inflation - Average of Point forecasts (SPF)

Table A10: Aggregate data

For the aggregate variables used in Figure-1, the unemployment rate is not seasonally adjusted and corresponds to the "Euro area-20 countries (from 2023)", as directly reported by Eurostat. The inflation rate refers to the headline HICP inflation for the "Euro area (EA11-1999, EA12-2001, EA13-2007, EA15-2008, EA16-2009, EA17-2011, EA18-2014, EA19-2015, EA20-2023)", reflecting the evolving composition of the Euro Area over time. Inflation expectations are one-year-ahead projections obtained from the Survey of Professional Forecasters (SPF), conducted by the ECB since 1999. The inflation gap shown in Figure 1 is defined as $\pi_t - \pi_t^e$, where π_t^e denotes the 12-month-ahead inflation expectations reported by professional forecasters.

E. Three-Equation Macro Model and Simulation

We consider the basic sticky price model augmented with internal dynamics, as developed in Bhattarai et al. (2014), and given below. Here, π_t denotes inflation, y_t denotes output, and R_t denotes the monetary policy rate.

$$\text{(1) Euler Equation:} \quad (Y_t - \eta Y_{t-1}) = (Y_{t+1} - \eta Y_t) \; - \; (1-\eta)\big(R_t - \pi_{t+1}\big) \; + \; e_{d,t}$$

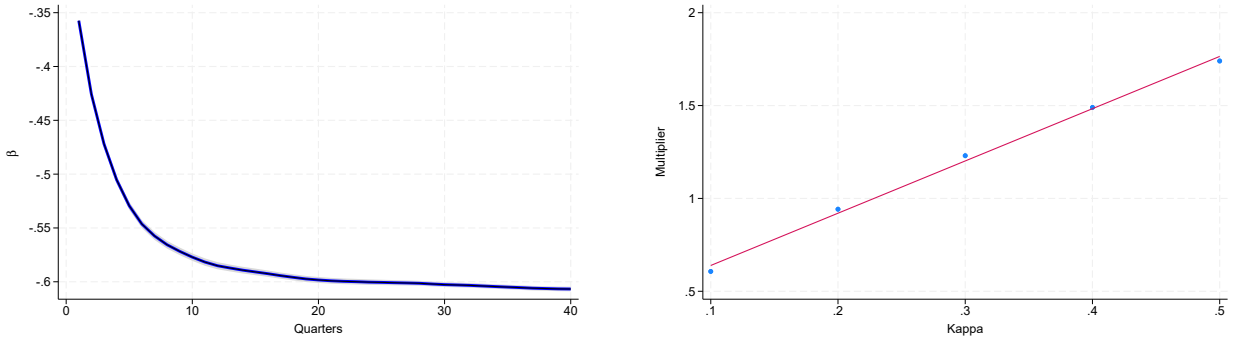
$$\text{(2) NK Phillips Curve:} \quad (\pi_t - \gamma \pi_{t-1}) = \beta(\pi_{t+1} - \gamma \pi_t) \; + \; \kappa \left[\varphi Y_t + \frac{Y_t - \eta Y_{t-1}}{1-\eta} \right] + e_{u,t}$$

$$\text{(3) Monetary Policy Rule:} \quad R_t = \rho_R R_{t-1} \; + \; (1-\rho_R)\big(\phi_\pi \pi_t + \phi_Y Y_t\big) + x_{R,t}, \quad x_{R,t} = \rho_{xR} x_{R,t-1} + \varepsilon_{R,t}$$

Table A11: Calibration

Parameter	Value	Description
η	0.00	Habit formation in consumption
β	0.99	Discount factor
γ	0.50	Price indexation
κ	0.10	Phillips-curve slope
φ	1.00	Inverse Frisch elasticity
ϕ_π	1.50	Taylor rule weight on inflation
ϕ_Y	0.00	Taylor rule weight on output
ρ_R	0.90	Interest-rate smoothing
ρ_{xR}	0.00	Policy-shock persistence
$\varepsilon_{R,t}$	0.01	Monetary policy shock

We simulate the model in the presence of the monetary policy shock, $\varepsilon_{R,t}$, and for the parameter values listed in the table. Figure A13a presents the first set of results, where it shows the multiplier responses (where we have flipped the sign to account for the negative relationship between output and unemployment). The multiplier response is dynamic due to internal dynamics in the model, and it converges after about 20 periods. These multiplier responses are analogous to our empirical results. Figure A13b then shows how the (medium-run) multiplier changes when we change the structural slope of the Phillips curve (κ) in the model. As is clear, the relationship is positive and monotonic.



(a) Benchmark Multiplier IRF

(b) Multiplier with Alternative κ

Figure A13: Phillips Multiplier For a Monetary Shock